

Introduction

Our understanding of mind and brain has come a long way over the millennia.

There was a time when people did not know that brain is the key organ responsible for our subjective experience. Greek philosopher Aristotle thought that the heart is the “seat of the soul” and the substrate for experience and selfhood. French philosopher Rene Descartes imagined that motor action is possible due to the action of “animal spirits” that rush through the nerves. The history of brain teaches us that though there was a considerable understanding of the brain’s structure (anatomy) even half a millennium ago, when it came to brain function, all sorts of fantasies were paraded as knowledge for many centuries.

Around the middle of 18th century, with developments in physics and physiology, a physics-based understanding of brain function began to take shape. It became clear that nerve signals are not “animal spirits” but electric signals not very different from the currents that flow in an electrical circuit. Developments in microscope revealed the peculiar hairy morphology of neurons, and presented a vision of brain as a network of neurons. Progress in neurochemistry, neuropharmacology and neurophysiology unraveled how neurons converse among themselves using chemical signals. Breakthroughs in technology is offering us vast treasures of neuroscientific data spanning many scales from single molecules, to neurons, to networks to whole brain and behavior.

Understanding the brain as an organ is very different from understanding other organs of the body. The brain, first and foremost, is an information processing machine. Like any other organ in the body, the brain too is a mass of cells. But unlike any other organ in the body, brain is a network of cells, a network that clearly distinguishes itself in its sheer size, complexity and lability. An adult brain has about 100 billion neurons, each with about 1,000-10,000 connections. We thus have a staggering figure of about 10^{14} - 10^{15} connections in the brain. Therefore, the brain network is perhaps more complex than the entire mobile network of the world, even if we assume that every one of the 7 billion denizens of the planet possess a mobile phone. Furthermore, this extremely complex cerebral network is quite labile, with neurons making and breaking connections at a time-scale that can be as short as a few tens of seconds.

Understanding brain function therefore means the ability to explain brain function in terms of the operations of this complex neural network. Thus the question “how does the brain see?” must be rephrased as “how do networks of neurons in the visual processing areas of the brain transduce the optical image that falls on the retina and process its many properties like form, color, motion etc?” And the answers to these questions are best clothed in the language of mathematics, which is the primary preoccupation of the science of computational neuroscience.

Models of brain can be classified broadly into two types: 1) biophysically realistic models and 2) abstract models. Biophysically realistic models are rooted in biophysics of neuron and brain, and aim to describe brain function in terms of electrical and chemical signaling of the brain. But these models can get extremely complex, computationally challenging, and often offer little insight into the essential information processing mechanisms that govern the function of a neural

system. Therefore, modelers constant try to strike a balance between neurobiological realism with reliable and convenient abstraction. The present course also follows a course of development that starts from a biophysically realistic description, tending towards more abstract models by progressive and systematic simplification.

The chapters of this course material are organized as follows.

Chapter 1 – presents elements of neurobiology that form the necessary preparation for a student of computational neuroscience. The first part of the chapter describes the biology of a single neuron, and the basic neuronal signaling mechanisms, electrical and chemical. The latter part describes the overall organization of human nervous system.

Chapter 2 – this chapter presents the mathematical ingredients that will be used in the subsequent chapters.

Chapter 3 – presents the Hodgkin-Huxley model of action potential generation. The Hodgkin-Huxley model is one of the first significant mathematical models of neuron function. It describes how the dynamics of neuronal ion channels generates a sharp spike in neuronal membrane potential known as action potential.

Chapter 4 – Although action potential generation is a key event in neural signaling, there are other important mechanisms, like dendritic processing, axonal propagation and synaptic transmission, which are described in the present chapter.

Chapter 5 – this chapter begins the process of simplification of the more complex, biophysical neural models described so far. Starting from FitzHugh-Nagumo neuron model, coursing via a series of progressively simplified models, it ends in the definition of the McCulloch-Pitts neuron model.

Chapter 5 marks the end of Part I of this course material. Armed with the simplified models presented at the end of Chapter 5, we describe abstract models in the subsequent chapters.

Chapter 6 – this chapter describes network models based on McCulloch-Pitts neuron model. The first of these is the perceptron which has only two layers of McCulloch-Pitts neurons. The second is the multilayer perceptron which is a generalization of the perceptron to arbitrary number of layers.

Chapter 7 – the multilayer perceptron belongs to a class of networks known as feedforward networks. The present chapter describes a model known as Hopfield network which has extensive feedback loops. The model serves as an associative memory. Ideas related associative memory are applied to describe the function of hippocampus towards the end of the chapter.

Chapter 8 – The previous chapter presented the idea of Hebbian learning, a learning mechanism used to store patterns in Hopfield network. The present chapter starts with Hebbian learning and shows how weights trained by Hebbian learning can extract principal components from input data. The chapter ends with Linsker’s model which uses Hebbian learning to explain evolution of response properties of neurons in the visual system.

Chapter 9 – This chapter introduces a variation of Hebbian learning known as competitive learning. The link between competitive learning and data clustering is explained. Next the chapter describes the Self-organizing Map (SOM), a neural network model based on competitive learning. The SOM model is applied to explain neural maps found in real brains.

Chapter 1

This chapter presents elements of neurobiology that form the necessary preparation for a student of computational neuroscience. The chapter is organized as follows. Section 1 describes the biology of a single neuron, and the basic neuronal signaling mechanisms, electrical and chemical. Section 2 describes the overall organization of human nervous system.

1.1 Neuron: Structure and Function

In this section, we give a brief outline of the structure and function of a neuron. At the outset, we must note that a neuron is not a specific cell, but a general term given to a large class of cells that share certain properties. This family of cells vary greatly in their morphology and the mechanisms and molecules they use to signal to each other. Therefore, it is impossible to describe the structure and function of neurons in all their rich variety in a short chapter. The present chapter gives a description of the structure and function of a typical neuron. The following chapter describes how the processes described verbally in the present chapter may be described mathematically.

1.1.1 Structure of a neuron:

- A neuron is primarily a cell. Therefore, like any other cell it has a cell body wrapped inside a cell membrane. It has a nucleus that contains the chromosomes which constitute the genetic information. It has other standard cellular components, the organelles like mitochondria, golgi bodies, nissil bodies, endoplasmic reticulum and so on. But what distinguishes a neuron from most other cells is the rich and elaborate wiring that seems to emerge from the cell body, also called *soma* (fig. 1.1.1.1). Neurons vary greatly in terms

of the number of these wiry structures that stick out of the cell body. While a neuron like the bipolar cell, found in the retina of the eye, has only two wires sticking out of the soma, there are neurons like the purkinje cell in the cerebellum, which have about one or two lakh wires per cell. In fact, even at a first glance, this wiring seems to be something odd about a neuron. One would not be totally off track if one surmised that it is these wires that make neuron a special cell, and, by extension, the brain a special organ. Hence broadly, depending on the Dendrite, we can classify the Neuron as (fig. 1.1.1.2).

- 1)Unipolar cell: 1 wire sticking out of the soma.
- 2)Bipolar cell: 2 wires sticking out of the soma.
- 3)Multipolar cell: Many wires sticking out of the soma.
- 4)Pseudounipolar Neuron

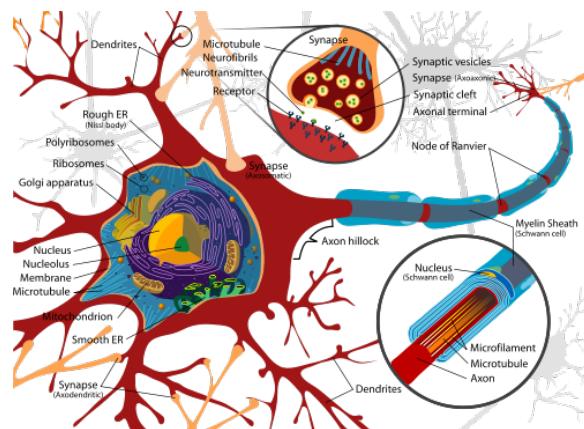


Figure 1.1.1.1: Internal structure of a neuron

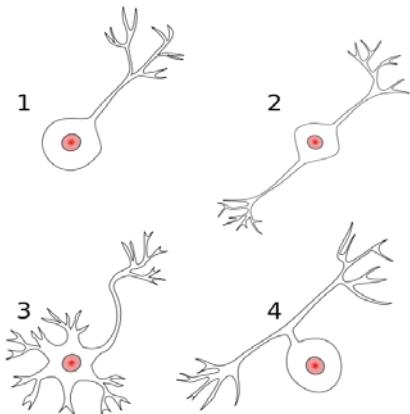


Figure 1.1.1.2: Classification of Neuron based on Dendrites

Structure of a Typical Neuron

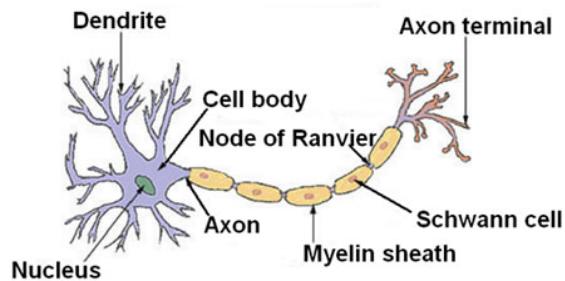


Figure 1.1.1.3: Structure of a neuron

On a closer look, one can distinguish two distinct portions in the wiring system: one portion has shorter, more densely distributed wiring, known as the dendrites; the other portion, consisting typically of a single long wire, known as an axon, branches out into smaller axon terminals at the far end. Neurons use these dendrites and axons to receive and transmit signals to each other. Signals *from* other neurons are received by the dendrites, while signals *to* other neurons are transmitted by axon and its terminals. Thus a neuron can be regarded as an input-output system

with dendrites as the inputs and the axon terminals as its outputs. Signals from one neuron to another are transmitted across a small gap – between the axon terminal of one neuron and the dendrite of another neuron – known as the synapse (fig. 1.1.1.3).

1.1.2 Electrophysiology of a neuron:

The basis of electrical signaling in a neuron is the fact that there is voltage difference between the interior of the neuron relative to the space surrounding the neuron, the extracellular space. This voltage, known as the membrane potential, is a constant (of about -70 mV) in resting conditions. However, when a neuron is stimulated, by one of several factors, the membrane voltage can show both positive and negative deviations. These voltage variations carry signals across the body of a neuron, and also contribute to signals that are transmitted from one neuron to another across the synapse.

Electrophysiology is a branch of physiology that deals with electrical phenomena related to biological systems. This field of science, in turn, has its roots in electrochemistry, a branch of chemistry that deals with the relationship between ions and electricity, a common example of which is a battery. A basic principle of electrochemistry states that when two compartments containing of an ion X, in different concentrations, are separated by a membrane that is selectively permeable to that ion, then, at equilibrium, voltage difference is generated between the two compartments. This voltage difference, known as the Nernst potential, depends on the ratio of equilibrium ion concentrations in the two compartments. Mathematical aspects of the Nernst potential are described in the next chapter.

1.1.3 Ions and ion channels:

Membrane potential in a neuron, or to that matter in any cell that has a membrane potential, is generated due to a difference in concentrations between the interior of the cell and the extracellular space. There is not one but several ions that contribute to the membrane potential, the key players being: Na^+ , K^+ , Cl^- and Ca^{2+} . The membrane potential is some sense, a combined effect, of the Nernst potentials of these individual ionic species. But to generate a Nernst potential we mentioned above that we require a semi-permeable membrane (fig. 1.1.3.1).. That is to generate a Nernst potential for Na^+ , we need the membrane to be semi- or selectively permeable to Na^+ , and so on. What about the membrane produces this semi-permeability or selective permeability?

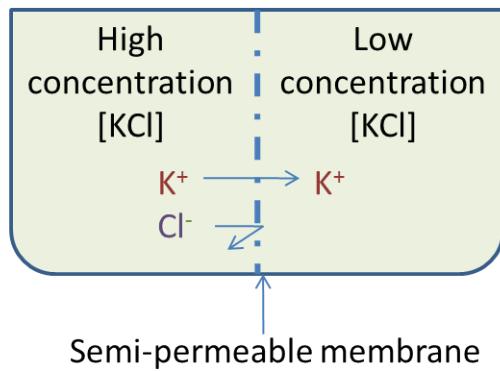


Figure 1.1.3.1: Neural Signaling with a semipermeable membrane

Neural membrane is impregnated with tiny pores, known as ion channels, that allow passage of ions. These channels, which are constituted by membrane spanning proteins, are usually selective to specific types of ions. Thus a sodium channel shows high selectivity to Na^+ ions,

though it allows passage of minute quantities of other ions. Similar is the case with channels of other ions – potassium channels, chloride channels, calcium channels etc.

Different ions are distributed differently across the neural membrane. For example, in resting conditions, sodium concentration is higher outside the cell, than inside. Therefore the Nernst potential of sodium is positive inside. Thus when only sodium ions are present, the membrane potential is equal to the Nernst potential of sodium, i.e., positive inside. On the contrary, potassium ion concentration is higher within the cell than without, making the corresponding Nernst potential negative inside. Thus when only potassium ions are present, Nernst potential is negative inside.

$$V_1 - V_2 = \frac{RT}{Z_x F} \ln \frac{[X]_2}{[X]_1}$$

V₁-V₂ - Nernst potential for ion 'X'

[X]_{1,2} – concentrations of 'X'

Z_x – Valence of 'X'

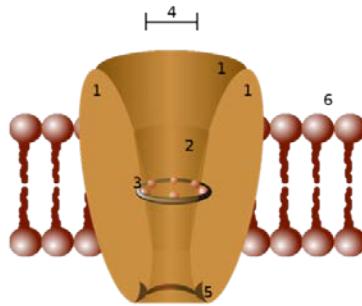
R – Ideal gas constant

T – Absolute temperature

F – Faradays' constant

RT/F = 26 mV at T=25°C (Z_x = +1)

The ability of a channel (fig. 1.1.3.2).to allow ions is not fixed for all time. Channels can be in OPEN or CLOSED states. A channel in OPEN state has naturally greater permeability than in CLOSED state. OPEN channels also offer greater electrical conductance to the ions to which they are permeable.



- **1** - channel domains(typically four per channel),
- **2** - outer vestibule,
- **3** - selectivity filter,
- **4** - diameter of selectivity filter,
- **5** - phosphorylation site, **6** - cell membrane.

Figure 1.1.3.2: Structure of a Channel

The conductance of a channel determines the contribution of the Nernst potential of the corresponding ionic species, to the membrane potential. Earlier we stated that, when only potassium ions are present, the membrane potential is negative, and when only sodium ions are present, the membrane potential is positive. Now we consider a small variation of that scenario. Consider a situation when both sodium and potassium ions are present (with the usual intra- and extra-cellular distributions) but only potassium channels are open (which is effectively the same as not having sodium ions). Again we have a negative membrane potential. Similarly if only sodium channels are open, we have a positive membrane potential. Thus we can think of sodium and potassium channels are knobs for turning the membrane potential up and down. To keep the membrane voltage negative, as it is in the resting conditions, keep the potassium channels open, with sodium channels closed. To take the membrane potentials to positive levels, open the sodium channels and shut down the potassium channels. These intuitive considerations will be made more rigorous mathematically, in the following chapter.

1.1.4 Ion channels and Gating:

We mentioned above that ion channels can be switched between OPEN and CLOSED states by various factors. This switching of the state of ion channels is known as *gating*.

There are 4 factors that drive channel gating: a) ligand-gating, b) voltage-gating, 3) phosphorylation c) stretch.

Ligand gating: In ligand-gating (fig. 1.1.4.1) a molecule binds to the ion channel and changes its confirmation (shape), thereby changing its open/close state.

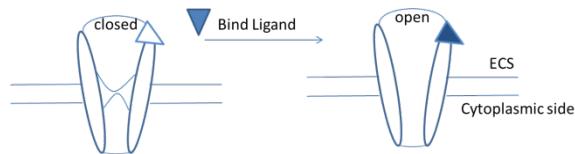


Figure 1.1.4.1: Ligand Gating

Voltage-gating: In this form of gating (fig. 1.1.7), it is the membrane voltage that controls channel gating. The fact that these channels are gated by voltage implies that their conductance is a function of membrane voltage. Therefore, the voltage-current characteristics of a voltage-gated or voltage-sensitive channel, are nonlinear. These channels do not obey the Ohm's law. These nonlinear conductances are crucial in generating neural signals in the form of sharp voltage spikes known as *action potentials*.

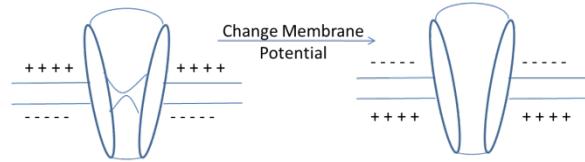


Figure 1.1.4.2: Voltage Gating

Phosphorylation-gating: In this form of gating (fig. 1.1.4.3), the channel goes to an open state from closed state when a phosphate group is attached to the channel, a process known as phosphorylation. The energy required for opening the channel comes from the phosphate group.

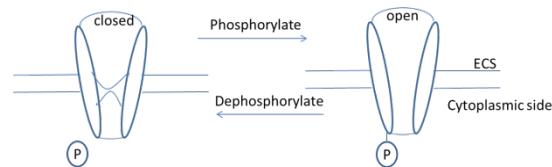


Figure 1.1.4.3: Phosphorylation Gating

Stretch-gating: Stretch of the cell membrane can also open channels. This mechanism converts a stretch event into an electrical event, since the opened channels permit current. Such channels (fig. 1.1.4.4) are found, for example, in nerve endings that transduce touch in skin.

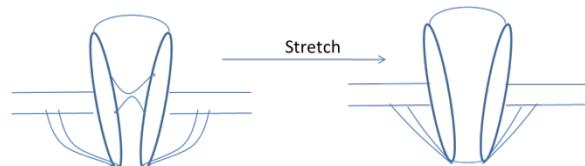


Figure 1.1.4.4: Stretch Gating

Each neuron receives signals via its dendrites, from other neurons; the dendritic signals flow towards the soma and combined at a part of the soma called the axon hillock; signals that arise from the axon hillock then propagate along the axon; at the end of the axon collaterals the axonal signals are transmitted to other neurons via the synapse. The cycle continues...

Thus, signaling in a neuron may be divided into four stages (fig. 1.1.5.1):

1.1.5 Stages of Neural Signaling:

- 1) Signaling over the dendritic tree
- 2) Summation at the axon hillock
- 3) Signal propagation along the axon and its collaterals
- 4) Signal transmission across the synapse

Neural signals are electrical and chemical:

So far we have been talking about neural “signals” rather vaguely without specifying the nature of these signals. Neural signals are electrical or chemical depending on the site of the signals. Among the 4 signaling components mentioned above, the first three are electrical, while the last one, signal transmission across the synapse, is a chemical step.

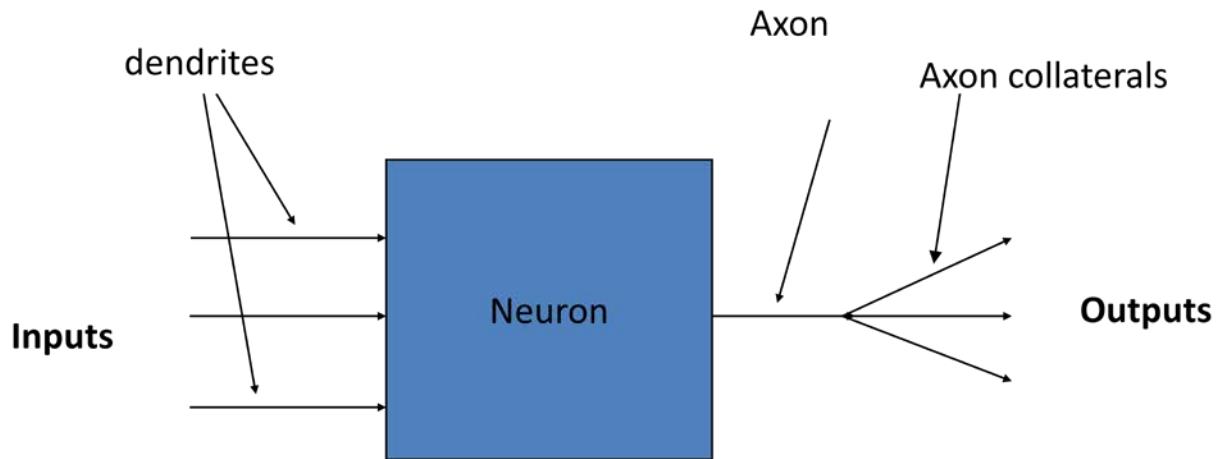


Figure 1.1.5.1: A neuron as an input/output system. Inputs from other neurons are received by the dendritic tree. Signals generated by a neuron are transmitted/broadcast to other neurons by the axon collaterals.

1.1.5.1 Dendritic signals:

The signals that are received at a dendritic branch from the axon terminal of another neuron, via a synapse, is in the form of a local voltage change. This voltage change can be either positive or negative, depending on the nature of the synapse. This voltage deviation propagates, like a wave, towards the cell body. Propagation along a dendrite is lossy (fig. 1.1.5.1.1). Therefore, the wave as it propagates loses amplitude and widens in time. Different waves, positive and negative, originating at different times, at different locations on the dendritic tree, flow towards the soma and get summated there.

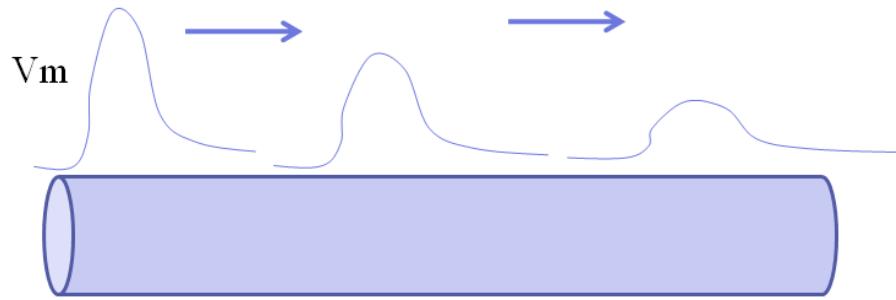


Figure 1.1.5.1.1: Dendritic propagation.

1.1.5.2 Summation at the soma and Action Potential:

The voltage waves arriving from the dendritic arbor, get summated at the soma (fig. 1.1.5.2.1), or more specifically the axon hillock, a knobby part of the soma at the root of the axon. The axon hillock has a high concentration of voltage-sensitive sodium and potassium channels, which is the reason behind a special response property of a neuron known as the “all-or-none” response. To illustrate this “all-or-none” response consider a thought-experiment in which you inject a pulse of current into a neuron. A small current pulse produces a small transient, upward deviation in the neuron voltage (fig. 1.1.5.2.2). As the current pulse amplitude is increased, amplitude of the voltage deviation also grows proportionally until the response reaches a threshold voltage level. Beyond this threshold, the neuron’s voltage continues to increase rapidly up to a high value and sharply drops to the original resting potential (fig. 1.1.5.2.3). This transient, rapid rise and fall of neuron voltage is known as Action Potential (AP). It has 3 phases:

- Rising Phase
- Falling Phase

- After Hyperpolarisation

Refractory period of the AP is a short period of time after an action potential has occurred, the cell membrane cannot fire another action potential. This can be *Absolute refractory period*: voltage-gated Na^+ channels won't open again or *Relative refractory period*: cell is hyperpolarized

Spatial Summation Temporal Summation
in Soma in Soma

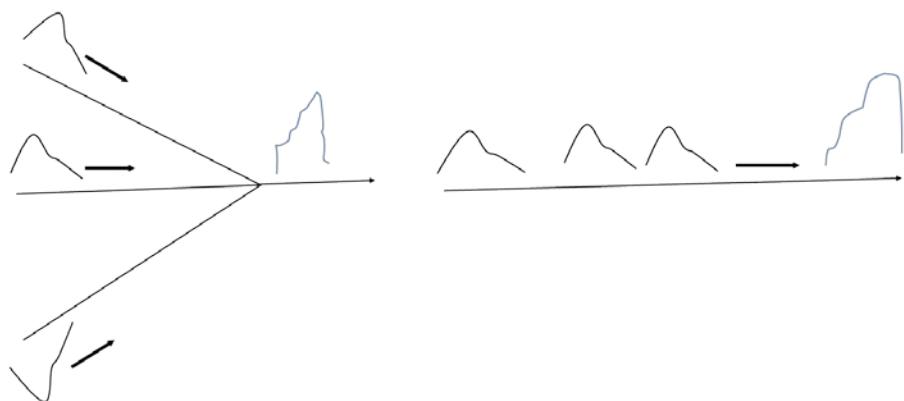


Figure 1.1.5.2.1: Summation at Soma

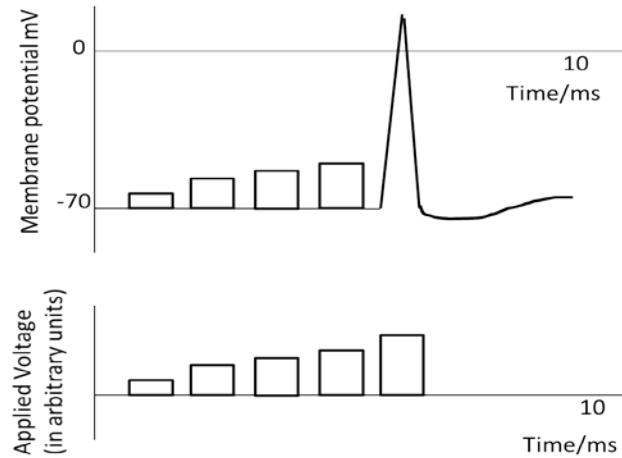


Figure 1.1.5.2.2: All or none Law

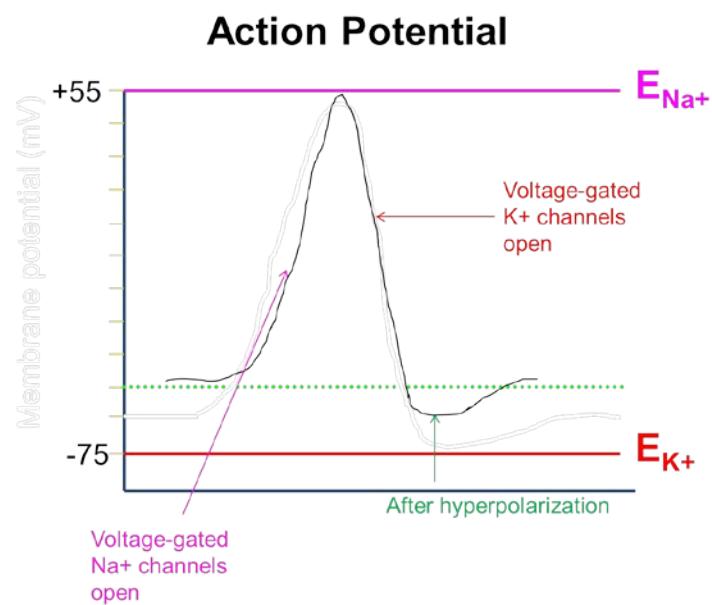


Figure 1.1.5.2.3: Phases in an Action Potential

When voltage waves of different amplitudes and signs flow towards the soma, they add up to change the local voltage of the soma. If that change is positive and exceeds a threshold value, an AP is produced at the axon hillock. If the net voltage change does not exceed the threshold value, no AP is produced (fig. 1.1.5.2.4). The state in which a neuron produces an AP is known as an *excited state*, as opposed to the usual resting state of the neuron when the voltage is a low, constant value.

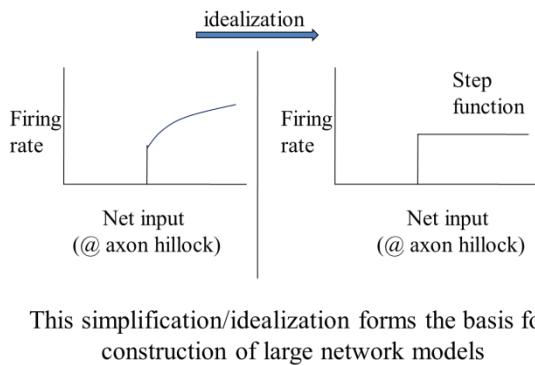


Figure 1.1.5.2.4: Neuron as a Thresholding Device

1.1.5.3 Axonal propagation:

The AP produced at the axon hillock propagates along the axon and reaches the axon collaterals. An important difference lies in the manner in which an AP propagates along the axon, and a voltage wave propagates along the dendrites. A voltage wave loses its amplitude and also spreads in time as it propagates down the dendritic tree towards the soma. But an AP propagates

intact, without losing amplitude or spreading in time, as it propagates along an axon (fig. 1.1.5.3.1).

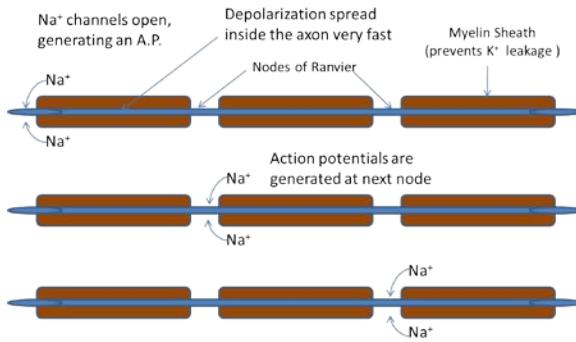


Figure 1.1.5.3.1: Axonal Propagation

This is possible because of presence of special molecular machinery on the axon and also in the axon hillock. This machinery is responsible for charging up the signal as it propagates along the cable (fig. 1.1.5.3.2).

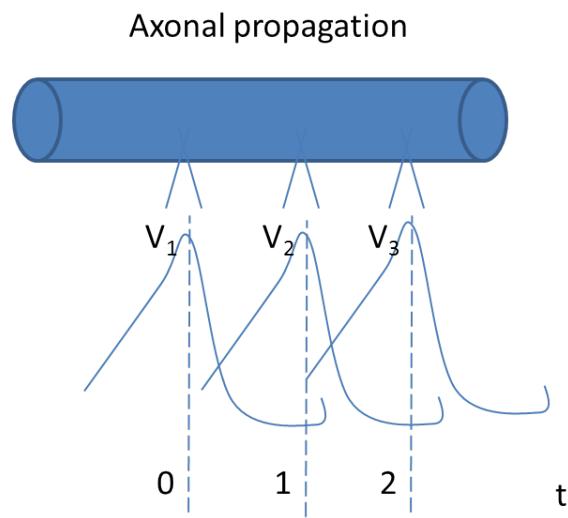


Figure 1.1.5.3.2: Axonal Propagation

1.1.5.4 **Synapse**

The synapse forms the functional connection between two neurons. It is a convenient window through which one neuron signals to another. The synapse is what makes the brain a *network* of neurons and not a mass of cells. Most importantly, it is the site of most learning and memory in the brain. It is now believed that whenever the brain learns something, the result of learning is somehow encoded in the properties of synapses. Recognition of the importance of the synapse in understanding brain function has led to a whole movement known as “Connectionism.”

The synapse is a special structure where all the neurochemical machinery for neuron to neuron signaling is concentrated. Its design ensures that a signal released by a neuron has maximal effect on a target neuron with minimum attenuation. Based on the nature of signal used, synapses are classified as i) electrical synapses and ii) chemical synapses.

1.1.5.4.1 **Electrical synapses**

Electrical synapses are direct cell-to-cell contacts. They are mediated by gap junctions, which form corridors that directly connect cytosols of two neurons (fig. 1.1.18). These corridors can permit passage of ions and small molecules. Thus two neurons coupled by gap junctions can electrically signal to each other by exchange of ions.

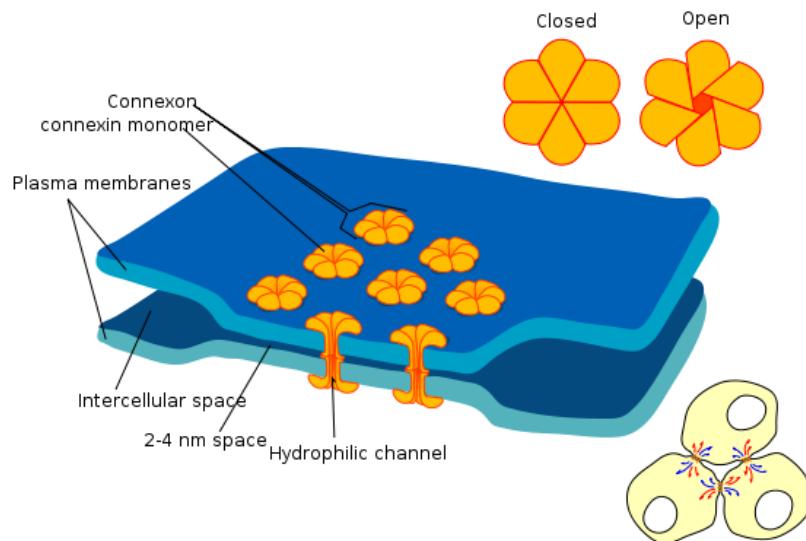


Figure 1.1.5.4.1: Electrical Synapse

Electrical synapses are typically bidirectional i.e., signal propagates in both directions (from neuron A to neuron B and back). This is because the electrical synapses may be simply regarded as a passive conductance that links the membrane voltages of two neurons. (There are, however, rectifying gap junctions in which conductance is greater in one direction than the other, analogous to a diode in electronic circuits. These may still be considered bidirectional with some asymmetry.)

1.1.5.4.2 Chemical synapses

In a chemical synapse, signaling is mediated by a chemical messenger that transduces electrical changes in one neuron and translates it into appropriate electrical changes in a target neuron. Fig. 1.1.5.4.2 shows the microstructure of a synapse.

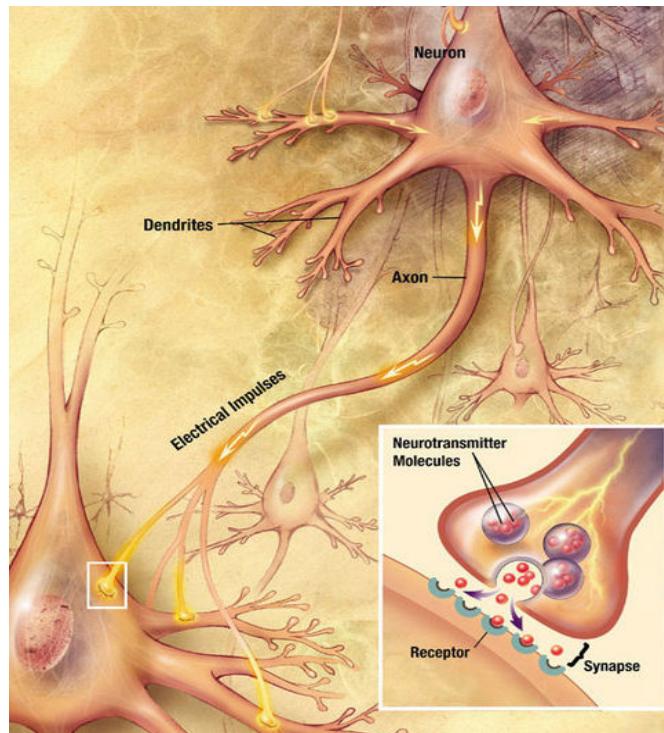


Figure 1.1.5.4.2: Chemical Synapse

Some terminology is in line. Transmission across a chemical synapse is unidirectional proceeding from the *presynaptic* neuron to the *postsynaptic* neuron. Synapse is simply a site where the axon terminal of the presynaptic neuron, known as the presynaptic terminal, and an appropriate part of the postsynaptic neuron, known as the postsynaptic terminal, are held in close proximity. The pre- and post- synaptic terminals have no physical contact and are separated by a gap known as the *synaptic cleft*. The cleft is rather narrow with a gap of only about 20 nm. Such close apposition of the pre- and post-synaptic terminals in a synapse makes possible a reliable transmission with minimal loss

1.1.6 Neurotransmission:

As mentioned above signaling processes up to the time when an AP arrives at an axon terminal are electrical, while the signal transmission across the synapse involves exchange of chemicals. When an AP arrives at an axon terminal, a special substance, known as the neurotransmitter, is released from the terminal, by exocytosis. This release is probabilistic and every AP need not produce a PSP. The neurotransmitter diffuses through a 20 nm gap contained in the synapse, known as the synaptic cleft, and reaches the dendrite of another neuron.

Action of neurotransmitter on the dendrite produces a voltage change in the dendrite known as the Post Synaptic Potential. This is simply the positive or negative deviation from the membrane potential of the dendrite, as described above. A positive voltage is more likely to propagate to soma and increase its local voltage, contributing its subsequent excitation. Therefore, a positive change is known as the Excitatory Post Synaptic Potential (EPSP). On the other hand, a negative change, when it propagates to the soma, tends to prevent the neuron from

excitation. In other words, it tends to inhibit the neuron. Therefore, such voltage change is known as the Inhibitory Post Synaptic Potential (IPSP). A synapse at which a EPSP is produced is called an excitatory synapse while the synapse at which an IPSP is produced is an inhibitory synapse. Now let us take a closer look at the process of neurotransmission. What makes one transmission event produce an EPSP as opposed to an IPSP? There are three molecular players in the process of neurotransmission: neurotransmitter, receptor and an ion channel. When the transmitter molecules released from the presynaptic terminal diffuse towards the postsynaptic terminal and bind with a receptor, which is associated with an ion channel. The binding event between the neurotransmitter and the receptor opens the associated ion channel. If the channel that is opened is a Na^+ channel, the local, postsynaptic membrane potential increases, which is the EPSP (fig. 1.1.6.1). On the other hand, when the ion channel that is opened is a K^+ or a Cl^- channel, it results in a negative deviation in the postsynaptic membrane voltage, which is an IPSP (fig. 1.1.6.2).

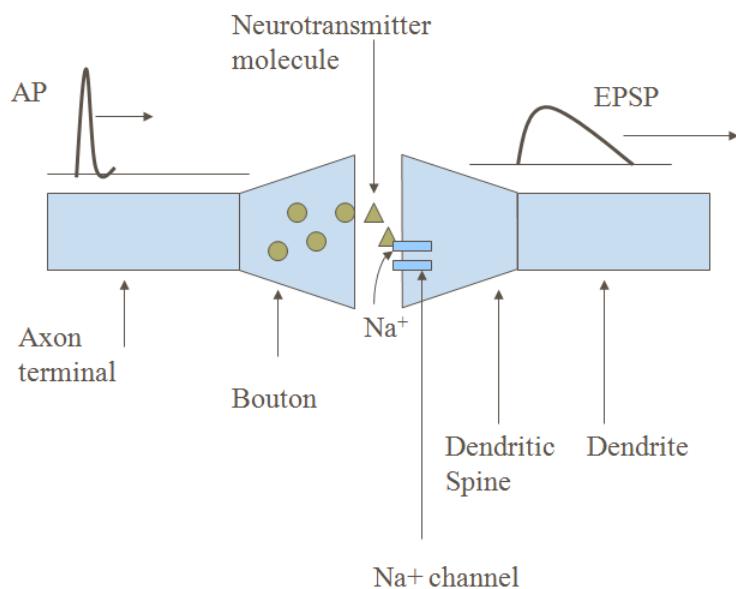


Figure 1.1.6.1: Excitatory Post Synaptic Potential (EPSP)

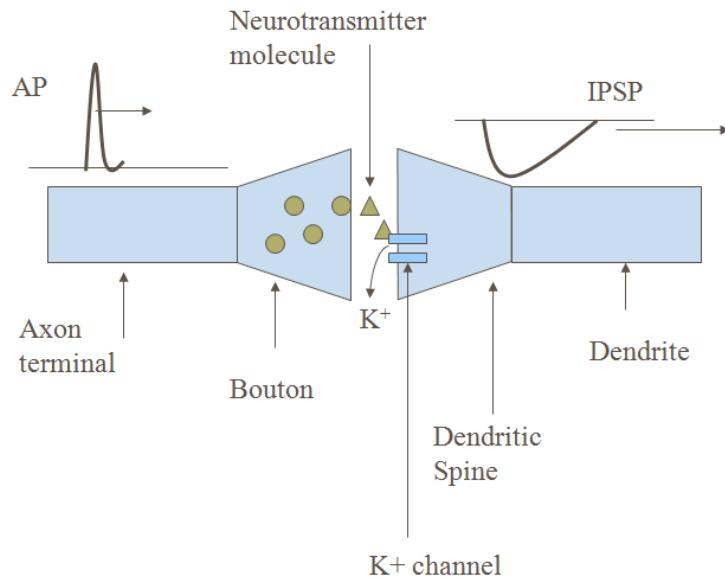


Figure 1.1.6.2: Inhibitory Post Synaptic Potential (IPSP)

Let us consider a few basic facts about neurotransmitter, receptor and ion channels.

1.1.6.1 Neurotransmitter:

In terms of their action, neurotransmitters may be broadly classified into excitatory and inhibitory neurotransmitters. A neurotransmitter that produces an EPSP is known as an excitatory neurotransmitter, an important example of which is glutamate. Similarly a neurotransmitter that produces an IPSP is an inhibitory neurotransmitter, a key example of which is Gamma Aminobutyric Acid (GABA). Glutamate and GABA are two important neurotransmitters in the brain. Over half of all brain synapses release glutamate, and about 30-40% of all brain synapses release GABA. Increased glutamate transmission tends to increase the overall excitation in the brain, while GABA

transmission has the opposite effect. The push-pull effect due to glutamate and GABA keep the brain in a state of balance between over-excitation and total inhibition.

Neurotransmitters vary in terms of their speed of action, scope of action, spatial extent of their action etc. However, in terms of chemistry, the classes of neurotransmitters: i) amino acids, ii) biogenic amines, iii) others and iv) neuropeptides.

The below three categories of neurotransmitter are grouped under the general class of *fast-acting neurotransmitters*, since their post-synaptic effects show up at the time scales of milliseconds.

- i) Amino Acids: There are 4 main candidates in this category: glutamate, aspartate, amino-butyric acid (GABA) and glycine. These are fast acting, capable of producing post-synaptic currents within a few milliseconds. Glutamate and aspartate are prominent excitatory transmitters, while GABA and glycine are inhibitory. Most rapid neurotransmission in vertebrate nervous systems is mediated by this class of neurotransmitters.
- ii) Biogenic Amines: There are 5 substances in this category: acetylcholine (Ach), norepinephrine, dopamine, serotonin and histamine. Their activity is much slower than the amino acid neurotransmitters, lasting over a duration of hundreds of milliseconds. However, as it is almost always the case in biology, there are exceptions. The speed of transmission depends on the type of receptor. For example, Acetylcholine acts fast when the transmission is mediate by a nicotinic receptor and slow acting in case of a muscarinic receptor.
- iii) Others: adenosine, nitric oxide etc.

In addition to the above 3 classes, there is another broad class of neurotransmitters known as neuropeptides. These are basically peptides, short chains of amino acids, which can have action

on the postsynaptic side. The neuropeptides are said to be slow-acting since their postsynaptic action occurs over time scales of seconds.

iv) Neuropeptides β -endorphin is an important example of a neuropeptide which interacts with opioid receptors in the brain. These are short chains of amino acids, acting over a time scale of minutes or more. This class of neurotransmitters is sometimes present along with (“colocalized”) with other fast-acting neurotransmitters. Release of this type of neurotransmitter occurs at a greater firing rate of the presynaptic stimulus, compared to what it takes to release fast acting neurotransmitters.

A question that often arises regarding neurotransmitters is: does a synapse release a single neurotransmitter or multiple neurotransmitters? A general principle of neurotransmission, called Dale’s principle, states that a neuron releases only a single transmitter. This is not true since a single neuron can release multiple transmitters from its synapses, a phenomenon known as *cotransmission*. Thus a modified form of Dale’s principle, due to Sir John Eccles, states that “at all the axonal branches of a neuron, there was liberation of the same transmitter substance or substances.” Thus the same set of neurotransmitters are released by a neuron throughout its lifetime at its synapses.

1.1.6.2 Receptors

Receptors are grouped as 1) ionotropic or 2) metabotropic receptors, based on the manner in which they interact with the associated ion channel.

- 1) In ionotropic receptors, the receptor or the binding site is located on another part of the protein complex that forms the ion channel.
- 2) In metabotropic receptors, the binding of neurotransmitter and receptor activates a series of intra-cellular events, one of which is the channel opening.

A given neurotransmitter can have several receptors.

Examples:

- 1) Glutamate receptors: 3 ionotropic and 1 metabotropic.
 - a. Ionotropic receptors: N-methyl D-aspartate (NMDA) receptor, α -amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA) receptor, kainate receptor.
 - b. Metabotropic receptor: mGluR receptor.
- 2) GABA receptors: are of two classes.
 - a. GABAa receptors are ionotropic
 - b. GABAb receptors are metabotropic

1.1.7 Structural Classification of Synapses:

Above we described a synapse structurally as a meeting point between axon terminal of the presynaptic neuron and an “appropriate part” of the post-synaptic neuron. Synapses are actually classified on the basis of what this “appropriate part” of the post-synaptic neuron is.

- 1) Axodendritic synapses: These synapses occur between axon terminals of one neuron and a dendrite of another. Further, this contact can occur at the shaft of the dendrite, or at a dendritic spine. It has been observed that synapses on dendritic spines are usually excitatory. When a signal is transmitted across an axodendritic synapse it undergoes

certain attenuation as it propagates down the dendritic tree towards the soma of the post-synaptic neuron. This propagation also involves a certain delay.

- 2) Axosomatic synapses: In these synapses, the axon terminal of one neuron makes contact with the soma of another. Typically these synapses happen to be inhibitory. By virtue of their proximity to the soma, signals across axosomatic synapses are more effective than signals from axodendritic synapses.
- 3) Axoaxonic synapses: Existence of axoaxonic synapses seems counterintuitive since a synapse is supposed to be junction between the output end of a neuron (its axon terminal) and the input end (dendrite or soma) of another neuron. However such synapses do exist. Signal across an axoaxonic synapse typically modulates transmission across one of the above two types of neurons.

1.1.8 Synaptic Transmission:

To summarize, a synapse simply converts an electrical event on the pre- side (arrival of AP) into an electrical event on the post- side (EPSP or IPSP). This conversion is done by an intermediate chemical step mediated by the trinity of Neurotransmitter-Receptor-Ion Channel. Fig. 1.1.8.1 shows the major structural components of a chemical synapse.

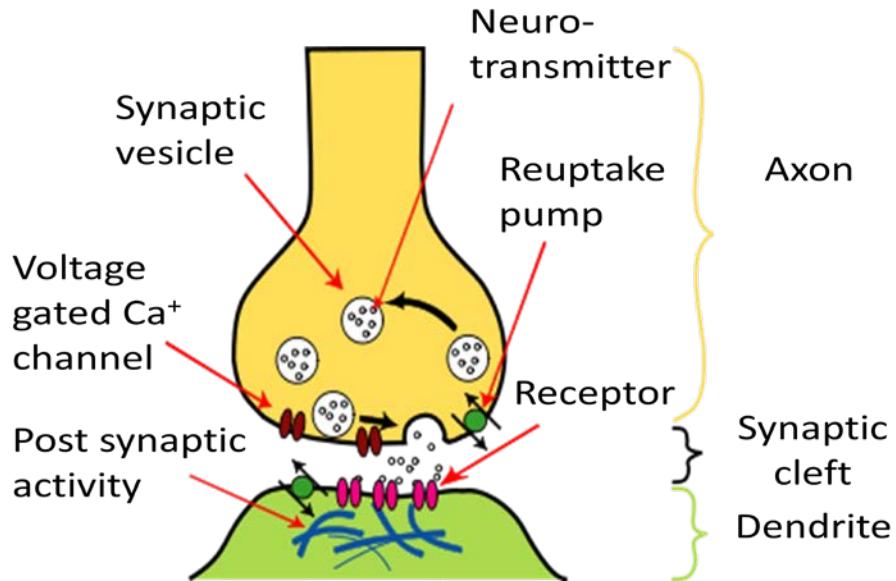


Figure 1.1.8.1: Major structural components of a synapse

How does this transformation of an AP on the pre- side to a PSP on the post- side take place? Complex molecular machinery is used for this purpose. Let us take a closer look at the series of events involved in this conversion.

Step 1: AP generated at the axon hillock arrives at the axon terminal or the pre-synaptic terminal.

Step 2: AP arrival increases local membrane potential in the pre-synaptic terminal.

Step 3: Voltage-sensitive Ca^{++} channels open. Ca^{++} rushes into the pre-synaptic terminal.

Step 4: Increased Ca^{++} concentration in the pre-synaptic terminal causes vesicles containing neurotransmitter to fuse with pre-synaptic membrane and release

neurotransmitter into the synaptic cleft. Such spewing out of material by cells is known as exocytosis.

Step 5: Neurotransmitter diffuses through the synaptic cleft and binds to receptor molecules on the post-synaptic membrane

Step 6: The binding event signals to associated ion channels to open/close

Step 7: Current influx/efflux through the open channels produces a (E/I)PSP across the post-synaptic membrane.

There are certain secondary, “clean up” events that accompany the above events.

Step 4b: (Vesicle recycling): When vesicles fuse with the pre-synaptic membrane, the membrane of the vesicle fuses with plasma membrane increasing the surface area of the pre-synaptic membrane. If this process goes on indefinitely, there will not be any membrane left for synthesis of new vesicles. However, the vesicular membrane patch that fused with the pre-synaptic membrane is recycled, taken back into the pre-synaptic terminal for synthesizing new vesicles. This process is thought to occur via several intricate mechanisms.

Step 5b: Transmitter molecules that are released into the synaptic cleft do not linger there for ever. They are removed from that region by several mechanisms. This removal is necessary to terminate the transmission process. Otherwise prolonged exposure of post-synaptic receptors to transmitter molecules makes the receptors insensitive to subsequent signals coming from the pre-synaptic side. Removal of transmitter from the cleft is done by 3 mechanisms: 1) diffusion, 2) enzymatic degradation (breaking down of transmitter by specific enzymes) and 3) re-uptake (transmitter molecules are actively taken back into pre-synaptic terminals and packaged into newly synthesized vesicles).

1.1.9 Synaptic Strength:

Since the PSP produced as a result of a single AP on the presynaptic side has variable magnitude, we can introduce the notion of synaptic strength. There is no single, textbook definition of synaptic strength. Biologically speaking there are several factors, both pre- and post-synaptic, that determine the synaptic strength. A reasonable definition of synaptic strength could be:

Synaptic strength = average PSP produced in response to an AP on the presynaptic side.

Synaptic strength figures as an important quantity in any discussion of information processing in the brain, since this strength varies as an effect of learning and memory. Results of learning and memory seem to be coded in the form of synaptic strength. This labile quality of synapses is known as synaptic plasticity.

1.2 Organization of the Nervous System: An Overview

The nervous system can be broadly classified into two major categories – the Central Nervous System (CNS) and Peripheral Nervous System (PNS) (Fig. 1.2.1).

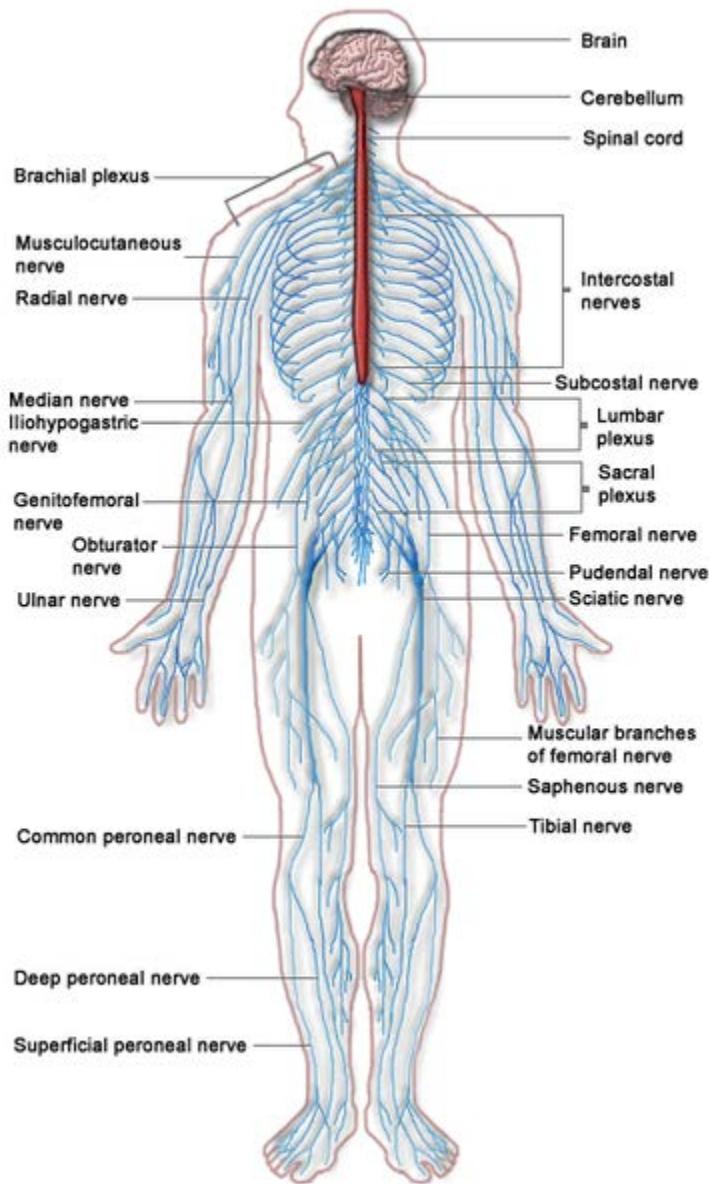


Figure 1.2.1: Human Nervous System

1.2.1 Central Nervous System

The central nervous system consists of the brain and spinal cord.

- **Brain**

The brain consists of the cerebrum, cerebellum and brain stem (containing the midbrain, pons and medulla oblongata).

- **Spinal Cord**

The **spinal cord** is a long, thin, tubular bundle of nervous tissue and support cells that extends from the brain (the medulla oblongata specifically). The brain and spinal cord together make up the central nervous system (CNS).

1.2.2 Peripheral Nervous System

The peripheral nervous system is divided into two parts – the somatic nervous system and the autonomic nervous system.

- **Somatic Nervous System**

The somatic nervous system performs two major functions – sensory and motor. The sensory nerves innervate skin, muscles and joints, and provide information about muscle and limb position, etc. It is majorly composed of 12 cranial nerves and 33 spinal nerves.

- **Autonomic Nervous System**

The autonomic nervous system can be classified into two major parts – the sympathetic and parasympathetic nervous systems. The sympathetic system participates in the body's reaction to stress, and helps in reacting to an emergency ("fight or flight") situation. The parasympathetic system conserves body resources

and maintains homeostasis. About 75% of all parasympathetic nerve fibres are in the vagus nerve.

Example – Sympathetic nerves control the sudden increase in heart rate during a stressful situation, whereas the parasympathetic nerves participate in lowering the heart rate during rest.

A third part of the autonomic nervous system, called the enteric system, specifically controls the smooth muscles in the intestine.

1.2.3 Nerves

A nerve is an enclosed, cable-like bundle of axons (the long, slender projections of neurons) in the peripheral nervous system. A nerve provides a common pathway for the electrochemical nerve impulses that are transmitted along each of the axons.

There are three major types of nerves –

- Afferent nerves conduct signals from sensory neurons to the central nervous system, for example from the mechanoreceptors in skin.
- Efferent nerves conduct signals from the central nervous system along motor neurons to their target muscles and glands.
- Mixed nerves contain both afferent and efferent axons, and thus conduct both incoming sensory information and outgoing muscle commands in the same bundle.

Nerve fibres may also be classified into various categories based on the speed at which they conduct electrical impulses.

- Type A – Large, myelinated fibres (~120 m/s conduction velocity)
- Type C – Small, unmyelinated fibres (~0.5 m/s)

Myelin sheath – A sheath which may cover nerve fibres; it helps in faster conduction of electrical impulses.

1.2.4 Central Nervous System – Details

1.2.4.1 Cerebrum

The cerebrum consists of two hemispheres, joined by a large bundle of axons called the corpus callosum. The hemispheres are also connected by a smaller bundle of fibres called the anterior commissure.

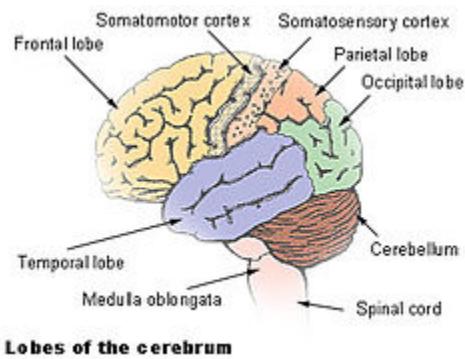


Figure 1.2.4.1.1: Lobes of the cerebrum.

The cerebrum can be broadly classified into four lobes –

- Frontal lobe - The **frontal lobe** is in the front of each cerebral hemisphere (Fig. 1.2.4.1.1 - 1.2.4.1.2). It is separated from the parietal lobe by a vertical gap called central sulcus, and from the temporal lobe by a deep fold called the lateral (Sylvian) sulcus. Primary motor cortex, which controls voluntary movements of the body is located in the

precentral gyrus, forming the posterior border of the frontal lobe.

Prefrontal cortex - The **prefrontal cortex** (P.F.C.) is the anterior part of the frontal lobes of the brain, lying in front of the motor and premotor areas.

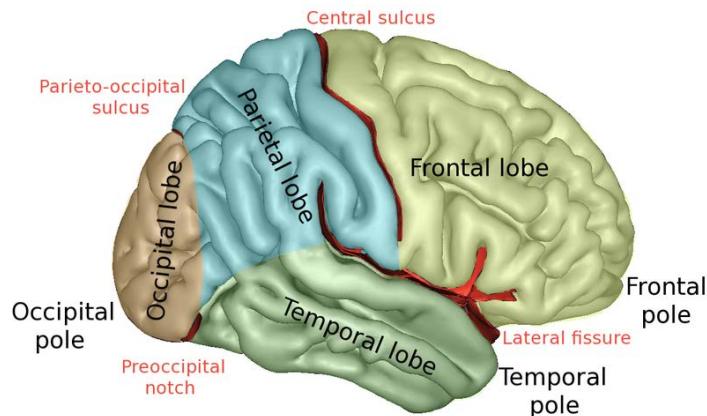


Figure 1.2.4.1.2: Lobes of the cerebrum.

- This brain region has been implicated in planning complex cognitive behavior, personality expression, decision making and moderating social behavior – a whole range of activities that are summarily described as executive function.
- Parietal Lobe - The **parietal lobe** is above the occipital lobe and behind the frontal lobe. It consists of the somatosensory cortex located at the anterior extreme, in the postcentral gyrus (Fig. 1.2.4.1.3). The somatosensory cortex processes touch information coming from mechanoreceptors located in the skin, muscles and joints.

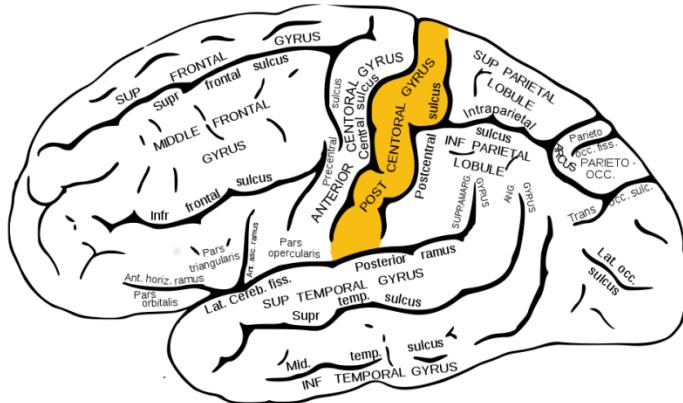


Figure 1.2.4.1.3: Post central Gyrus

Inferior parietal lobe consists of areas that process higher aspects of vision like spatial sense.

By virtue of its strategic location amidst three primary sensory cortices – somatosensory cortex, visual cortex and auditory cortex, - inferior parietal lobe as sensory association areas which integrate information from the three sensory modalities and extract abstract concepts.

- Temporal Lobe – The **temporal lobe** is located beneath the Sylvian fissure, a shared border between the temporal and frontal lobes.

Primary auditory cortex, involved in auditory processing, is located in the superior part of temporal lobe, bordering on the sylvian fissure.

The temporal lobe contains a deep structure called hippocampus, whose functions include spatial navigation, declarative memory, and memory consolidation.

Wernicke's area, which spans the region between temporal and parietal lobes, plays a key role, in language understanding, while Broca's area, which is in the frontal lobe, is responsible for language production (Fig. 1.2.4.1.4).

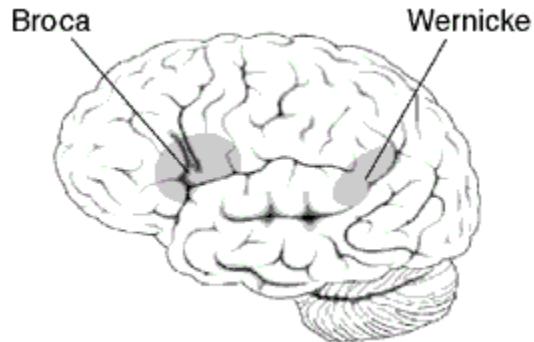


Figure 1.2.4.1.4: Lobes of the cerebrum.

The inferior part of temporal lobe has cortical areas that are responsible for recognizing complex visual objects, like for example, faces.

- Occipital Lobe – The **occipital lobe** has primary and higher areas of the visual cortex.
 - The primary visual cortex is commonly called V1 (visual one). It is partly located in the medial side of occipital lobe and partly in the posterior pole of the occipital lobe. V1 is often also called striate cortex because it can be identified by a large stripe of myelin, the Stria of Gennari. Visually driven regions outside V1 are called extrastriate cortex. There are many extrastriate regions, and these are specialized for different visual tasks, such as visuospatial processing, color discrimination and motion perception (Fig. 1.2.4.1.5).

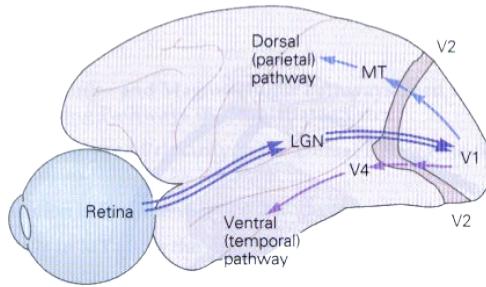


Figure 1.2.4.1.5: The Striate cortex

- Secondary and higher visual cortical areas extend from the occipital lobe and are spread over the neighboring parietal lobe.

Deep Brain Structures:

- Hypothalamus - The **hypothalamus** is an important structure controlling our autonomous function. It links the nervous system to the endocrine system via the pituitary gland. The hypothalamus controls body temperature, hunger, thirst, fatigue, sleep, and circadian cycles (Fig. 1.2.4.1.6).

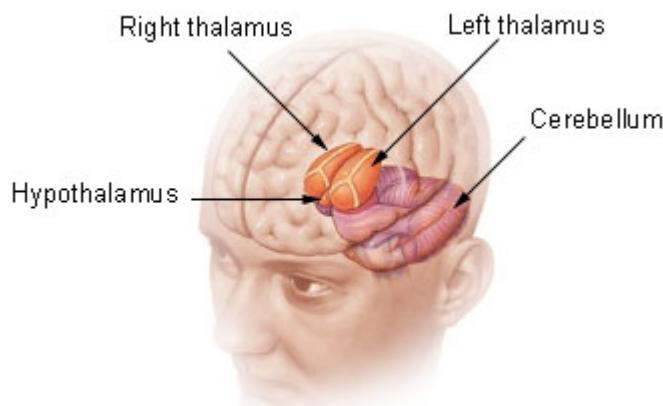


Figure 1.2.4.1.6: Thalamus and Hypothalamus

- Thalamus – The **thalamus** is located between the cerebral cortex and midbrain. It is an important hub through which most sensory information reaches the sensory cortex. It is involved in regulation of consciousness, sleep, and alertness (Fig. 1.2.4.1.6).
- Reticular Activating System - The **reticular activating system (RAS)** is an area of the brain responsible for regulating arousal and sleep-wake transitions.
- Limbic System - The **limbic system** is a set of brain structures involved in processing emotions, and is closely related to autonomous function. This system includes the hippocampus, amygdala, anterior thalamic nuclei, septum, limbic cortex and fornix (Fig. 1.2.4.1.7).

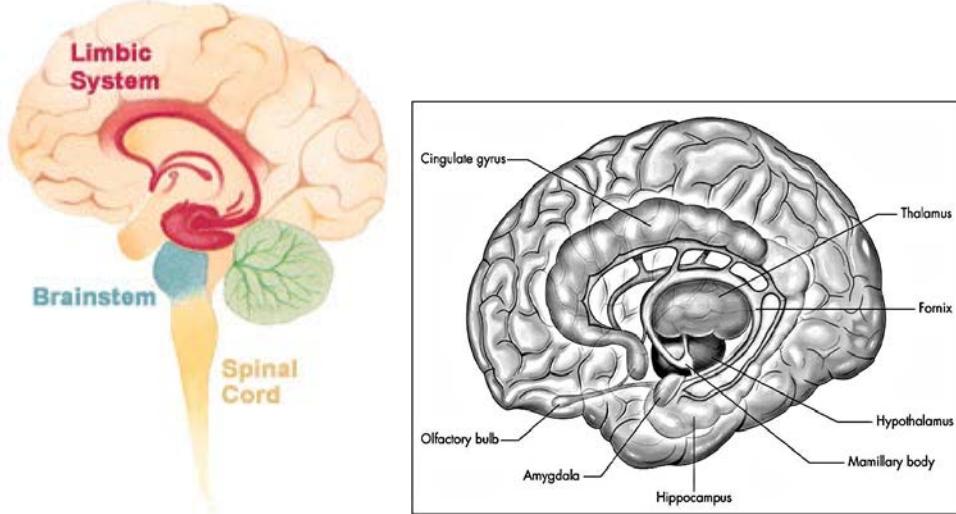


Figure 1.2.4.1.7: Limbic system

- Amygdala – This almond-shaped structure is located bilaterally in the medial temporal lobes of the brain. A part of the limbic system, this structure is involved in fear conditioning (Fig. 1.2.4.1.8).
- Basal Ganglia - The **basal ganglia** (or **basal nuclei**) are a deep brain circuit consisting of 6 or 7 nuclei. This circuit receives inputs from the cortex and projects back to the cortex. It has key functions like reward processing, action selection, working memory etc (Fig. 1.2.4.1.8)

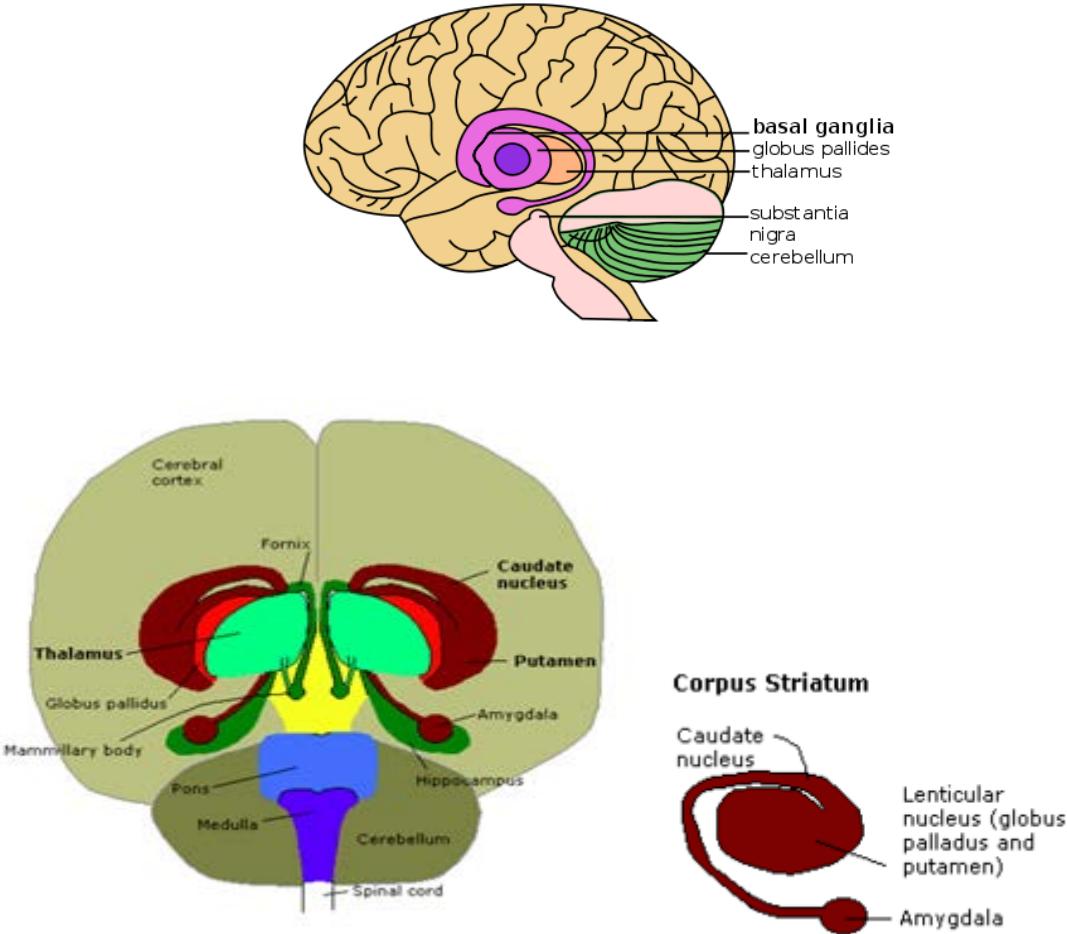


Figure 1.2.4.1.8: Diencephalon and Rhombencephalon

Damage to the basal ganglia can be associated to neurological and neuropsychiatric disorders like, for example, Parkinson's disease

The following section will discuss the brain stem and cerebellum (Fig. 1.2.4.1.8)–

- Brain Stem – The brain stem is a sort of a bridge area between the cerebrum and the spinal cord. It comprises of three regions: midbrain, medulla oblongata and pons. It consists of a lot of important control centers and anatomical features. Centers for

controlling cardiac and respiratory function; for regulating sleep and arousal; for controlling eating and drinking are located here.

- Medulla – This structure is a direct rostral (“towards nose”) extension of the spinal cord. Its primary functions include: regulation of blood pressure and respiration.
- Pons - The **pons** is a part of the brain stem, located below the midbrain and above medulla oblongata. It consists of wiring that carries: 1) motor signals from the cerebrum down to the cerebellum and medulla, and 2) sensory signals from the body into the thalamus.
- Cerebellum - The **cerebellum** (Latin for *little brain*) is a distinct structure located posteriorly under the hemispheres. It is generally thought of as a motor control unit, though its role in cognitive and affective function is well-established. It is involved in control of fine movement, equilibrium, posture, and motor learning.
- Spinal Cord - The **spinal cord** is a long, tail-like bundle of nerve fibers that extends from the brain (the medulla oblongata specifically) along the spine. It has three key functions:
 - Carrying motor signals from the brain to the muscle
 - Carrying sensory information in the reverse direction
 - Coordinating reflexes

Chapter 2: Mathematical Preliminaries

2.1 Eigenvalue & Eigenvectors

1) A is a square matrix

$$A \in R^{n \times n} \quad (1)$$

$$AX = \lambda X \quad X \in C^n \text{ (n - Dimensional complex vector)}$$

$$\det(A - \lambda I) = 0$$

Let A be a real symmetric matrix

$$A^T = A \quad (2)$$

Or

Hermitian matrix

$$A^* = A \quad (3)$$

Theorem 1:

The Eigen values of a hermitian matrix λ_j are real.

Proof:

$$AX = \lambda X$$

$$X^T AX = \lambda X^T X$$

$$X^T AX = \lambda \|X\|^2$$

$$X^* AX = \lambda^* \|X\|^2$$

$$\lambda \|X\|^2 = \lambda^* \|X\|^2$$

λ is real

2) A is skew-symmetric

$$A^T = -A \quad (4)$$

Or

$$A^* = -A$$

$$3) \quad AX = \lambda X \quad (5)$$

$$X^T AX = \lambda X^T X$$

$$X^T AX = \lambda \|X\|^2$$

Transposing both sides,

$$X^T A^T X = \lambda^* \|X\|^2$$

$$-X^T AX = \lambda^* \|X\|^2$$

$$\lambda \|X\|^2 = -\lambda^* \|X\|^2$$

$$\lambda = -\lambda^*$$

λ is purely imaginary.

4) Q is an orthogonal matrix.

$$Q^T Q = Q Q^T = I \quad (6)$$

Then,

$$|\lambda| = 1 \quad (7)$$

Proof:

$$QX = \lambda X$$

$$X^T Q^T = \lambda^* X^T$$

$$X^T Q^T Q X = \lambda \lambda^* X X^T$$

$$X^T X = \lambda \lambda^* X X^T$$

$$\|X\|^2 = |\lambda| \|X\|^2$$

$$|\lambda|^2 = 1$$

$$\lambda = e^{i\theta}$$

5) For every symmetric matrix, A, there exists an orthogonal matrix, Q, such that

$$Q^T A Q = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n) \quad (8)$$

Where, $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ are eigenvalues of A.

Proof:

Let q_i be the eigenvector corresponding to λ_i

$$A q_i = \lambda_i q_i$$

Let D be a matrix such that,

$$D_{ij} = q_j A q_i$$

$$D_{ij} = q_j \lambda_i q_i$$

$$\text{if } \begin{cases} i = j & \lambda_i \\ \text{otherwise} & 0 \end{cases}$$

$$Q = [q_1, q_2, q_3, \dots, q_n]_{n \times n}$$

$$Q^T A Q = D$$

$$A = Q D Q^T$$

$A = \sum_{i=1}^n \lambda_i q_i q_i^T$ is called the eigen-decomposition of A.

6. Orthogonal transformation preserves lengths:

Proof:

$$\begin{aligned} Y &= QX \\ Y^T Y &= X^T Q^T QX \\ \|Y\|^2 &= \|X\|^2 \end{aligned} \tag{9}$$

Length of X equals length of Y.

2.2 Quadratic forms

2.2.1 Quadratic forms: Explanation

A quadratic form is a quadratic function associated with a symmetric matrix, A, and is expressed as:

$$E = X^T AX \tag{10}$$

A discussion of quadratic forms is relevant to study of neural systems since often it is required to minimize an “error function” that depends on a large number of parameters. To be able to do so, we must first define the minimum of a multivariate function. For a univariate function, f, the minimum is simply a point where $f'=0$ and $f''>0$. To define the minimum of a multivariate function, the concepts of f' and f'' must be generalized to multivariate functions.

For a multivariate function:

First derivative is the Gradient., $G = \nabla f$

Second derivative is the Hessian, H, is defined as,

$$H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} \tag{11}$$

Since H is a matrix and not a scalar like f'' , we need to specify how to define a minimum in terms of the Hessian H. It is here that we need quadratic forms. To understand this let us expand a multivariate function as a Taylor series:

$$f(X_0 + h) = f(X_0) + h^T G + \frac{1}{2!} h^T H h + \dots \text{(higher order terms)} \quad (12)$$

$f(X_0)$ is a constant and therefore does not affect the shape of $f(X)$ around X_0 .

Since $G = 0$, we may ignore the linear term $h^T G$.

The higher order terms beyond quadratic are too small and may be ignored.

Therefore, the function $f(X)$ has a minimum at X_0 , if the function $\frac{1}{2!} h^T H h$

has a minimum at $h = 0$. To verify this, we need to examine the shape of a quadratic function

$X^T A X$ around the origin ($X=0$).

It is easier to examine the shape of a quadratic function, in transformed coordinates, Y , obtained by rotating the original coordinates X , so that in the new coordinates the “cross-terms” are eliminated.

2.2.2 Geometric interpretation of eigenvalues in terms of quadratic forms:

For a real, symmetric matrix, A , let $Q^T A Q = D$,

(13)

where D is a diagonal matrix of eigenvalues.

Let $X = Qy$

$$E = y^T Q^T A Q y = y^T D y$$

$$E = \sum_i y_i^2 \lambda_i$$

Or,

$$\frac{\partial^2 E}{\partial y_i^2} = 2\lambda_i \quad (14)$$

Thus in the rotated coordinates, eigenvalues of A are proportional to the second derivatives of E .

Maximize the quadratic form, E , subject to $\|X\|=1$

This can be done using the method of Lagrangian multipliers as follows. Since there is only one constraint, the Lagrange function may be written as:

$$E' = \frac{1}{2} X^T A X - \lambda(X^T X - 1)$$

$$\nabla E' = 0$$

$$\nabla E' = AX - \lambda X$$

$$AX = \lambda X$$

$$X = q_i$$

That is, the directions along which E is stationary (derivative is 0) under the constraint of $\|X\|=1$,

correspond to the eigenvectors of the matrix A , to which the quadratic form is associated.

Thus we can see that the shape of a quadratic form at the origin depends on the composition of the eigenvalues of A . Here we come to the notion of definiteness. There are 5 kinds of definiteness.

- Positive definite :** An ‘ $n \times n$ ’ matrix ‘ A ’ is said to be positive definite if it meets the following condition:

$$\frac{1}{2} X^T A X > 0, \forall X \neq 0, X \in R^n \quad (15)$$

This implies $\lambda_i > 0$.

This case corresponds to the “minimum.” Thus a multivariate function has a minimum at a point X_0 , if the Hessian of the function at X_0 , is positive definite.

- Positive semi-definite:** An ‘ $n \times n$ ’ matrix ‘ A ’ is said to be positive semi-definite if it meets the following condition:

$$\frac{1}{2} X^T A X \geq 0, \forall X \neq 0, X \in R^n \quad (16)$$

This implies $\lambda_i \geq 0$.

- Negative definite:** An ‘ $n \times n$ ’ matrix ‘ A ’ is said to be negative definite if it meets the following condition:

$$\frac{1}{2} X^T A X < 0, \forall X \neq 0, X \in R^n \quad (17)$$

This implies $\lambda_i < 0$.

This case corresponds to the “maximum.” Thus a multivariate function has a minimum at a point X_0 , if the Hessian of the function at X_0 , is negative definite.

- 4. Negative semi-definite :** An ‘ $n \times n$ ’ matrix ‘A’ is said to be negative semi-definite if it meets the following condition:

$$\frac{1}{2}X^TAX \leq 0, \forall X \neq 0, X \in R^n \quad (18)$$

This implies $\lambda_i \leq 0$.

- 5. Indefinite:** An ‘ $n \times n$ ’ matrix ‘A’ is said to indefinite if it does not meet any of the above four conditions.

Illustrations of various kinds of definiteness for two-variable functions.

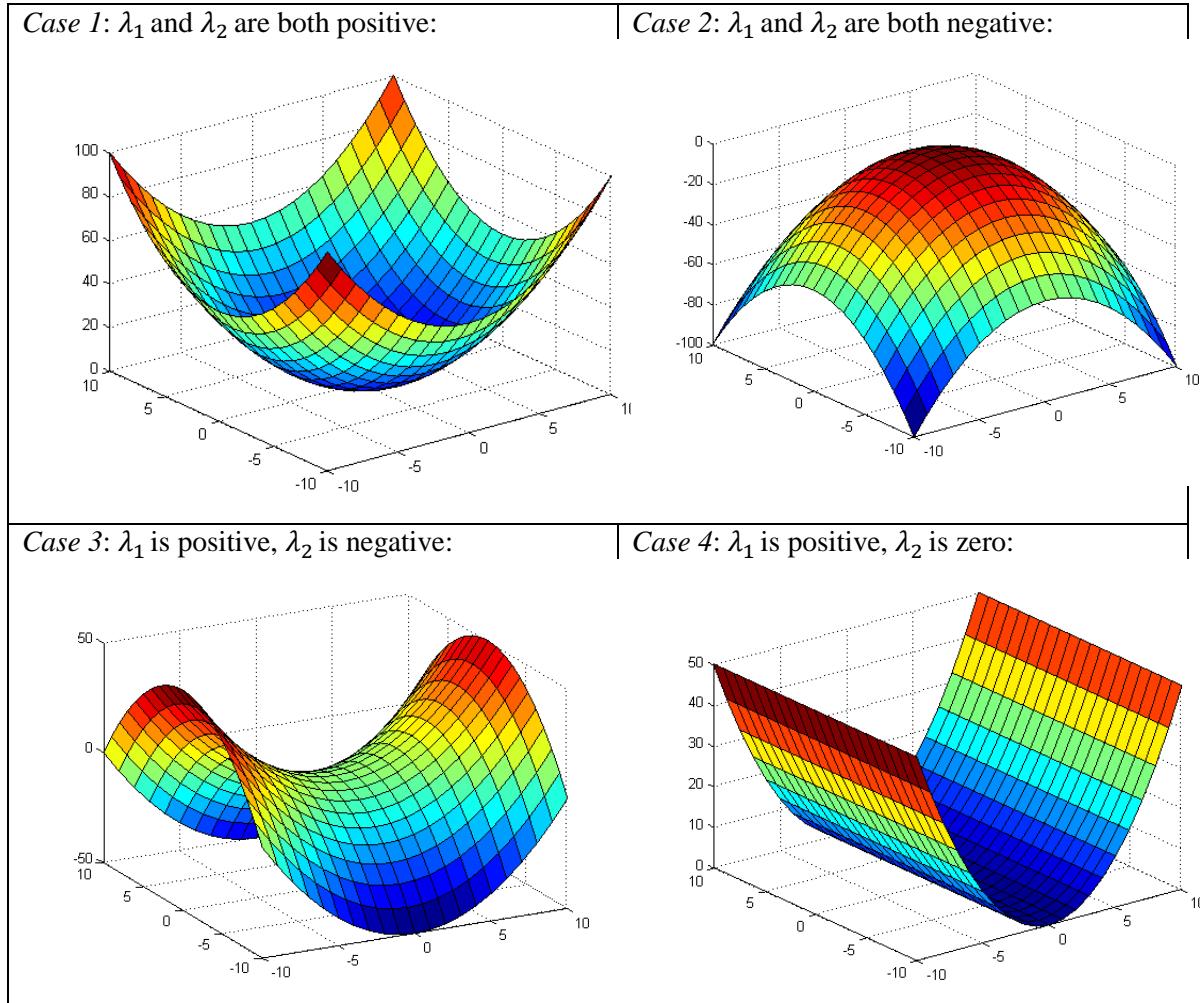


Figure 1: Example: For $n=2, X \in R^2, z = \frac{1}{2}(\lambda_1y_1^2 + \lambda_2y_2^2)$, plot z for all possible values of λ

Solution of linear equations: $AX=b$

Where, $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, $X \in \mathbb{R}^m$

Case 1

If $n=m$ (Unique solution)

Let A^{-1} exist.

$$X = A^{-1}B \quad (19)$$

Case 2(Least squares solution)

If $n>m$ (Under-determined case: more equations than unknowns)

Since there may not be a perfect solution, we attempt a Least Squares solution, by minimizing E .

$E = (b - AX)^T(b - AX)$. Minimizing E gives the following solution,

$$X_{LS} = (A^T A)^{-1} A^T b = A^+ b, \quad (20)$$

Where A^+ is called the pseudo inverse. Since $b = AX_{LS} = A(A^T A)^{-1} A^T b = AA^+b$

Case 3

$m>n$ (Infinite solutions)

Add minimum norm condition.

$$\text{Min } (\|x\|^2) \text{ such that } Ax = b. \quad (21)$$

Example Problem:

$E = x^2 + y^2$, constraint $ax + by = c$.

Let us use the method of Lagrangian Multipliers.

$$E' = x^2 + y^2 + \lambda (ax + by - c)$$

$$\frac{\partial E'}{\partial x} = 2x + a\lambda = 0$$

$$\frac{\partial E'}{\partial y} = 2y + b\lambda = 0$$

$$x = -a\lambda/2, y = -b\lambda/2$$

$$\lambda = \frac{-2c}{a^2 + b^2}$$

$$x = \frac{ac}{a^2 + b^2}$$

$$y = \frac{bc}{a^2 + b^2}$$

$$E = X^T X$$

$$E' = X^T X - \Lambda^T (AX - b)$$

$$\delta E' = 2X - (\Lambda^T A)^T = 0$$

$$X = A^T \Lambda / 2$$

Since $AX = b$

$$\frac{1}{2}AA^T\Lambda - b = 0$$

$$\Lambda = (AA^T)^{-1}b$$

$$X = A^T (AA^T)^{-1}b$$

2.3 Dynamic Systems and fixed points:

2.3.1 Dynamical Systems

Concepts from dynamic systems are most essential to study neural models. Ideas of stability, attractors, limit cycles, chaos appear again and again in discussions of brain dynamics. A few examples:

1. Memories in the brain are often modeled as *attractors* of brain dynamics.
2. The resting state of a neuron is considered as *stable node*, since the neuron returns to that state on small perturbations.
3. Periodic spiking activity of a neuron is modeled as a limit cycle.

A general dynamic system is defined as:

$$\frac{dx}{dt} = f(x), \quad x \in R^n \quad (22)$$

Thus x is a vector having n components, $x = [x_1, x_2, x_3, x_4, x_5, \dots]$

Stationary Points (also called, critical points, equilibrium points etc) are those where:

$$\frac{dx}{dt} = 0 = f(x), \quad \text{at } x = x_s \quad (23)$$

The behavior of the system around x_s , depends on the Jacobian of $f(x)$,

$$A_{ij} = \frac{\partial f_i(x)}{\partial x_j} \quad \text{Matrix of the partial derivatives is Jacobian} \quad (24)$$

If you linearize (22) around $x = x_s$,

$$\frac{dx}{dt} = Ax \quad (25)$$

$$\dot{x} = Ax$$

$$\dot{x}_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots$$

$$\dot{x}_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots$$

$$\dot{x}_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots$$

.

.

Hence

$$\dot{x}_n = a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots \dots \dots \dots + a_{nn}x_n$$

Whose solution would be,

$$X(t) = c_1 e^{\lambda_1 t} q_1 + c_2 e^{\lambda_2 t} q_2 + \dots \quad (26)$$

Verifying:

$$\dot{x}(t) = \sum c_i \lambda_i e^{\lambda_i t} q_i$$

$$= \sum c_i e^{\lambda_i t} A q_i$$

$$= A \sum c_i e^{\lambda_i t} q_i$$

$$= A x(t)$$

Types of stationary points ($n = 2$):

Description in terms of eigenvalues of A ,

(Plots generated in Matlab using quiver function):

1. Both eigenvalues are real, -ve (Fig. 2)

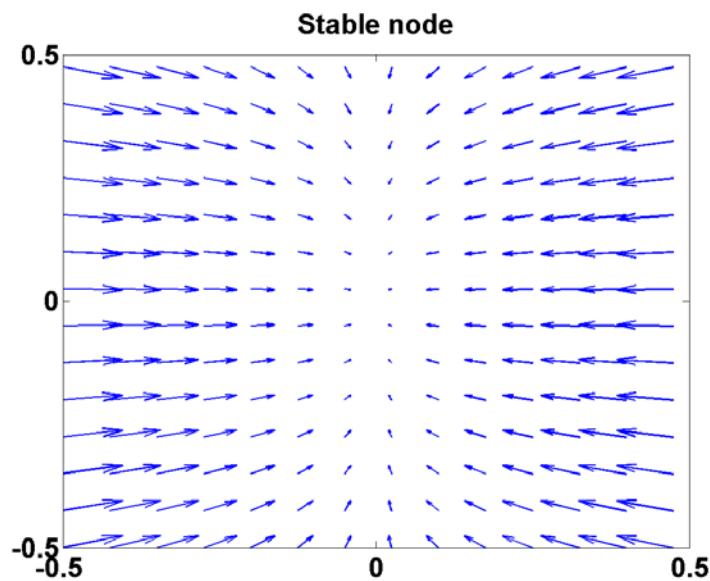


Figure 2: Stable node

2. Both eigenvalues real, +ve.

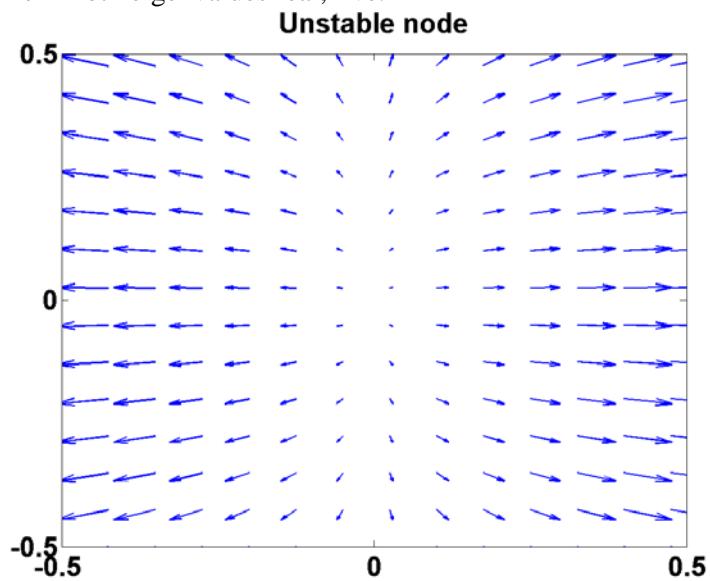


Figure 3: Unstable node

3. One +ve eigenvalue, one -ve eigenvalue

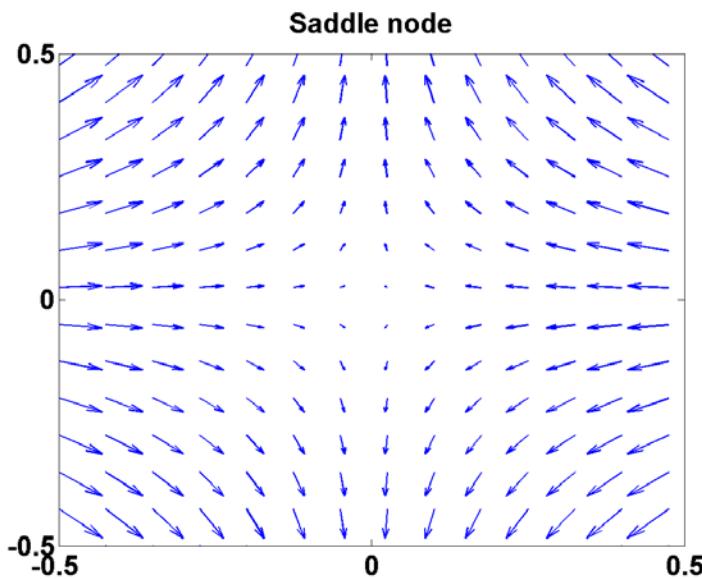


Figure 4: Saddle node

4. Both eigenvalues complex with -ve real parts.

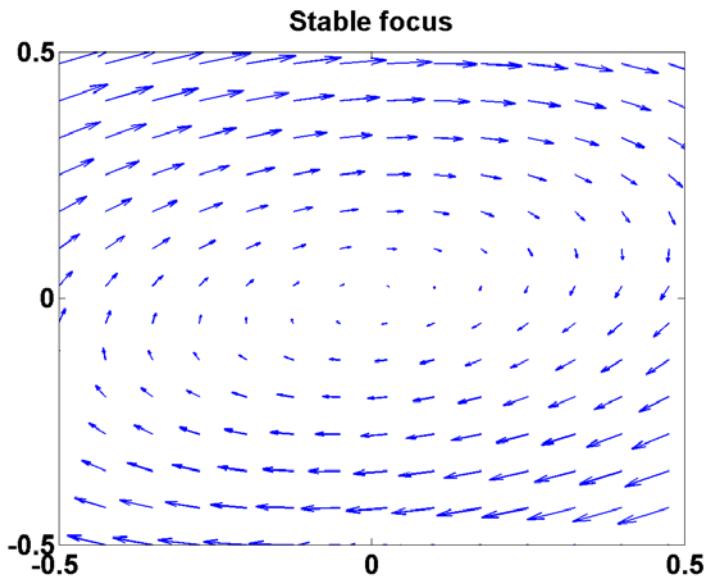


Figure 5: Stable focus

5. Both eigenvalues complex, +ve real parts.

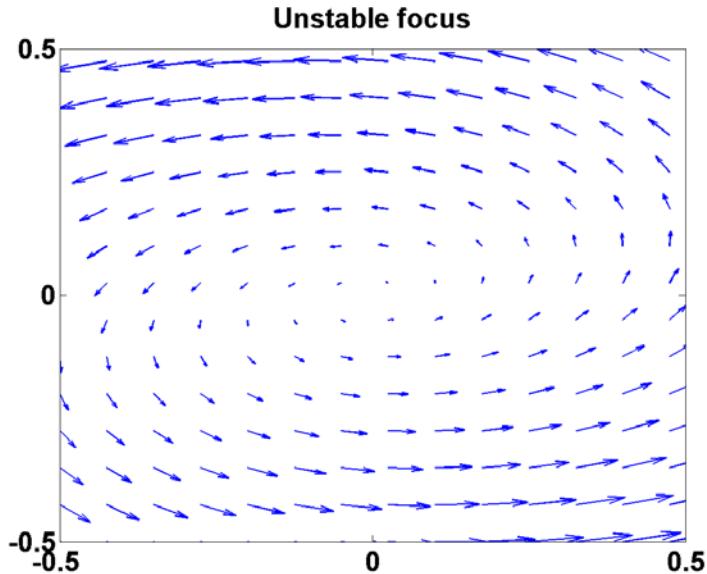


Figure 6: Unstable focus

6. Purely imaginary.

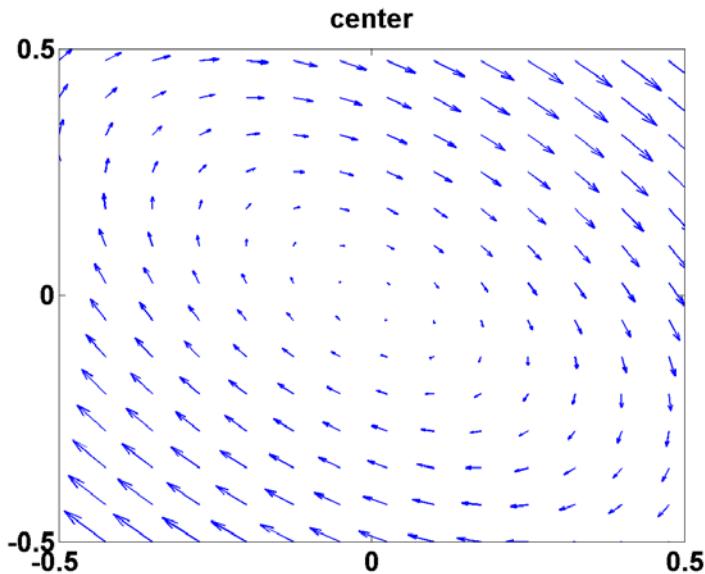


Figure 7: Center

Other cases:

7. Star: Both the eigenvalues being real, equal, negative

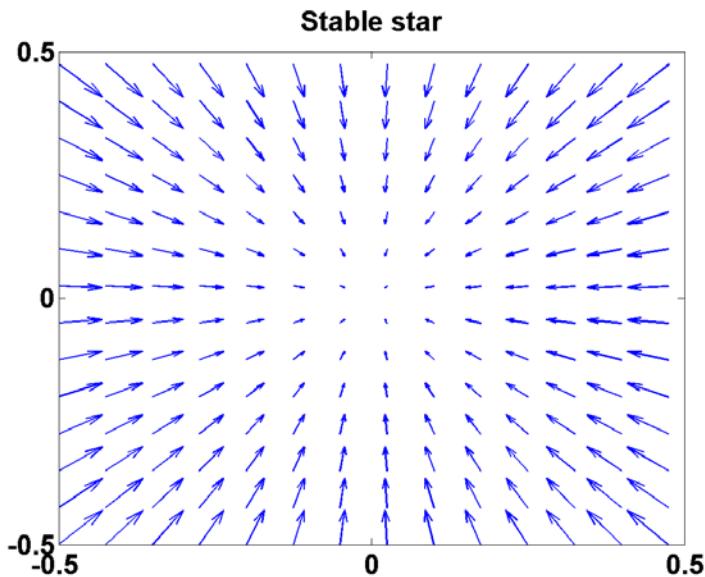


Figure 8: Stable star

8. One of the eigenvalues is zero.

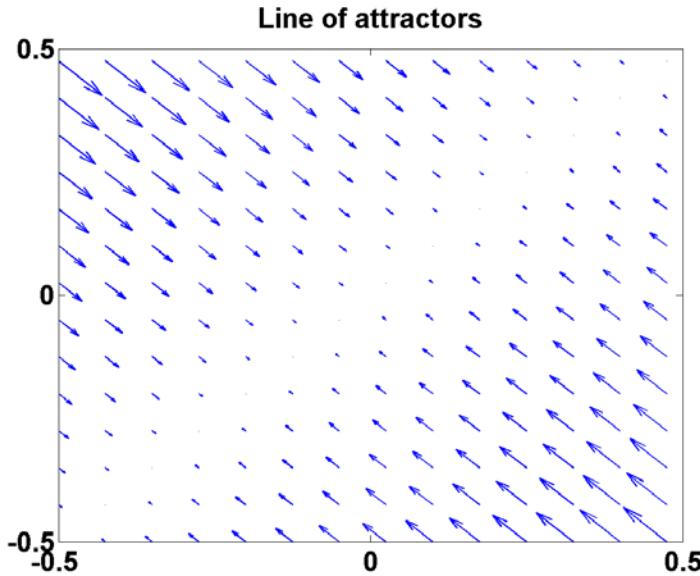


Figure 9: Line of attractors

2.3.2 Classification of Fixed Points:

For a linear dynamic system,

$$\dot{\mathbf{x}} = \mathbf{Ax} \quad (27)$$

the type of fixed point at the origin can be related to the trace and determinant of \mathbf{A} .

This relationship can be easily derived when \mathbf{A} is a 2×2 matrix, but the result applies to the general $n \times n$ matrix.

Let

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \text{ The characteristic equation is,}$$

$$\det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = 0, \quad (28)$$

Expanding the determinant, we get the following quadratic equation

$$\lambda^2 - \tau\lambda + \Delta = 0 \quad (30)$$

Where

$$\tau \equiv \text{trace}(A) = a + d$$

$$\Delta \equiv \det(A) = ad - bc$$

Solving for λ , we have,

$$\lambda_1, \lambda_2 = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2} \quad (31)$$

From the above form of the expression for the eigenvalues, λ_1, λ_2 , of the Jacobian of the dynamic system, $\dot{x} = Ax$, at the origin, we can infer a few things.

Case $\lambda < 0$:

Eigenvalues are real and have opposite signs. Therefore, the fixed point is a saddle node.

Case $\lambda > 0$:

If $\tau^2 - 4\Delta > 0$, both the roots are real. Further, $\sqrt{\tau^2 - 4\Delta}$ is less than $|\tau|$. Therefore,

- if τ is positive, both the roots are positive \rightarrow unstable nodes.
- if τ is negative, both the roots are negative \rightarrow stable nodes.

If $\tau^2 - 4\Delta < 0$, the roots are complex conjugates. Furthermore,

- If τ is positive, the real part of the roots is positive \rightarrow unstable focus
- If τ is negative, the real part of the roots is negative \rightarrow stable focus
- If τ is 0, the roots are purely imaginary \rightarrow center

If $\tau^2 - 4\Delta = 0$, the roots are equal \rightarrow line of fixed points.

The above scheme of classification is summarized in the “map” shown below.

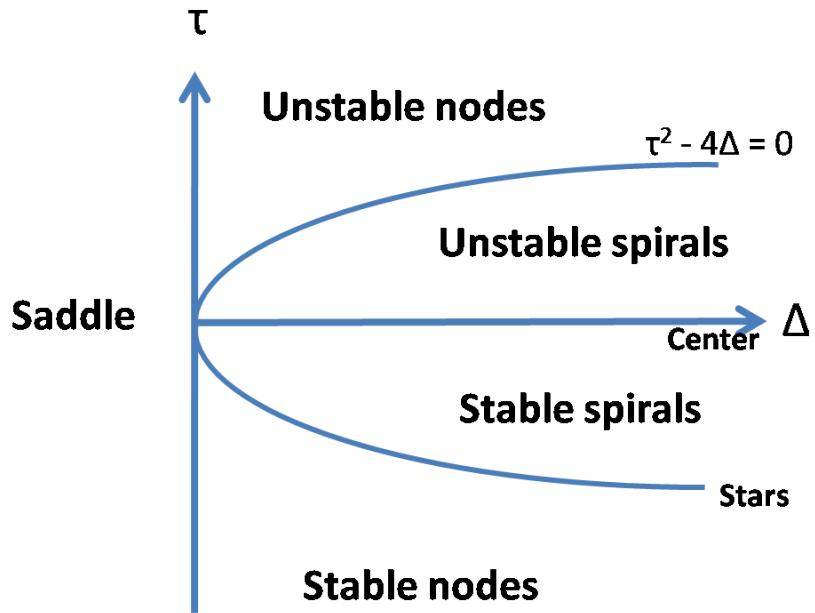


Figure 10: Map of Δ vs τ

2.3.3 Phase-plane Analysis:

Above we have briefly listed out the types of fixed points that occur in a linear n-dimensional dynamic system. The origin is the only fixed point in such systems. But a nonlinear dynamic system can have multiple fixed points with a much larger repertoire of behaviors, and therefore much harder to analyze than a linear system. Therefore it is convenient to study two-dimensional nonlinear systems which offer several advantages:

- They can be easily visualized
- They have rich dynamics and are sufficiently interesting study (for example, they exhibit limit cycle behavior, which can be used to model rhythmic behavior in real systems)
- They have a limited range of dynamics compared to n-dimensional ($n > 2$) nonlinear dynamic systems. (for example, two-dimensional continuous, differentiable dynamic systems do not exhibit chaos).

For these reasons, two-dimensional systems occupy a special place in study of dynamic systems.

A general form of a two-dimensional dynamic system can be expressed as:

$$\dot{x} = f(x, y) \quad (32)$$

$$\dot{y} = g(x, y) \quad (33)$$

The simplest kind of analysis that can be performed on such a system is to identify all the fixed points and classify each of them by local analysis.

Fixed points of eqns. (32-33) may be calculated by solving,

$$\dot{x} = f(x, y) = 0 \quad (34)$$

$$\dot{y} = g(x, y) = 0 \quad (35)$$

Eqns (34-35) represent curves in the x-y plane and are known as null-clines.

$f(x, y) = 0$ is the x-nullcline and $g(x, y) = 0$ is the y-nullcline. The intersection points of the two null-clines are the fixed points of the system.

Let (x_0, y_0) is a fixed point of the system described by eqns. (34-35). Dynamics in the neighborhood of the fixed point may be expressed as:

$$\dot{\varepsilon}_x = f(x_0, y_0) + \varepsilon_x \frac{\partial f}{\partial x} + \varepsilon_y \frac{\partial f}{\partial y} + \text{higher order terms} \quad (36)$$

$$\dot{\varepsilon}_y = g(x_0, y_0) + \varepsilon_x \frac{\partial g}{\partial x} + \varepsilon_y \frac{\partial g}{\partial y} + \text{higher order terms} \quad (37)$$

A linear approximation of the above system is given as,

$$\begin{bmatrix} \dot{\varepsilon}_x \\ \dot{\varepsilon}_y \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}}_J \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \end{bmatrix} \quad (38)$$

Where J denotes the Jacobian of the system eqns. (32-33)

The above equation may be expressed more compactly as,

$$\dot{\varepsilon} = J\varepsilon$$

The fixed point at $\varepsilon = 0$, can be classified by performing eigenvalue analysis of the above 2D system.

2.3.4 Limit Cycles:

Limit cycles are a type of oscillatory behavior characterized by two properties:

- Periodicity:

A system exhibiting limit cycle behavior is confined to a closed loop trajectory and periodically visits every point on that loop. Thus, if a system whose state variable is denoted by $x(t)$, visits a point, a , at time, t , and if ‘ a ’ is on a limit cycle, then

$$x(t) = x(t+T)=a, \quad (39)$$

where T is the period of the limit cycle.

- Isolatedness

This property refers to the behavior of the system when it is slightly perturbed from the limit cycle. When a system begins at a point that is in the neighborhood of a limit cycle, it either approaches the limit cycle asymptotically, or moves away from it. In this respect, a limit cycle is different from a ‘center.’ In case of a center, when the system is slightly perturbed, it continues on a new periodic orbit and does not return to the previous orbit.

Limit cycles are relevant to neural dynamics because neural spiking activity may be conveniently modeled as a limit cycle.

There are several ways in which it can be proved that a system exhibits limit cycle behavior, some of which are discussed below.

- 1) Special systems
- 2) Lienard Systems
- 3) Poincare-Benedixson Theorem

Explanation:

- 1) **Special systems:** There is no general rule for determining if a system has limit cycle behavior. This absence of general, universal methods is characteristic of nonlinear systems. But some systems have a convenient form so that it can be easily shown that they have limit cycle behavior.

Example: We can easily show the system given below has limit cycles –

$$\dot{x} = -y + \mu x(1 - x^2 - y^2) \quad (40)$$

$$\dot{y} = x + \mu y(1 - x^2 - y^2) \quad (41)$$

Let us rewrite the above equations in polar form by substituting:

$$x = r \cos(\theta) \text{ and} \quad (42)$$

$$y = r \sin(\theta) \quad (43)$$

Combining eqns. (40-41) as shown below

$$x\dot{x} + y\dot{y} = \mu(x^2 + y^2)(1 - x^2 - y^2) \quad (44)$$

and expressing the result in polar coordinates we have

$$d(r^2)/dt = \mu r^2(1 - r^2) \quad (45)$$

Or,

$$\dot{r} = \frac{\mu}{2} r(1 - r^2) \quad (46)$$

Now consider,

$$\theta = \arctan(y/x) \quad (47)$$

Differentiating both sides,

$$\dot{\theta} = \frac{1}{1 + \frac{y^2}{x^2}} \frac{x\dot{y} - y\dot{x}}{x^2} = \frac{x\dot{y} - y\dot{x}}{r^2} \quad (48)$$

Using eqns. (40-41) we have,

$$x\dot{y} - y\dot{x} = x^2 + y^2 = r^2 \quad (49)$$

Combining eqn. (48) and eqn. (49) we have,

$$\dot{\theta} = 1 \quad (50)$$

Thus eqns. (40-41) are re-expressed as, eqns. (46) and eqn. (50).

From eqn. (46) we can see that r approaches the stable value of 1: dr/dt is positive for $r < 1$, and negative for $r > 1$.

Thus we have a limit cycle which is a circle of unit radius and angular velocity of 1.

- 2) **Lienard systems:** This represents a slight improvement over showing limit cycles only in individual systems. Lienard systems are a general class of systems that exhibit limit cycles under certain conditions.

Liénard's equation is equivalent to the system

$$\dot{x} = y \quad (51)$$

$$\dot{y} = -g(x) - f(x)y. \quad (52)$$

The following theorem states that this system has a unique, stable limit cycle under appropriate hypotheses on f and g . For a proof, see Jordan and Smith (1987), Grimshaw (1990), or Perko (1991).

Liénard's Theorem: Suppose that $f(x)$ and $g(x)$ satisfy the following conditions :

- (1) $f(x)$ and $g(x)$ are continuously differentiable for all x ;
- (2) $g(-x) = -g(x)$ for all x (i.e., $g(x)$ is an odd function) ;
- (3) $g(x) > 0$ for $x > 0$;
- (4) $f(-x) = f(x)$ for all x (i.e., $f(x)$ is an even function) ;
- (5) The odd function $F(x) = \int_0^x f(u) du$ has exactly one positive zero at $x = a$, is negative for $0 < x < a$, is positive and nondecreasing for $x > a$, and $F(x) \rightarrow \infty$ as $x \rightarrow \infty$.

Then the system (51-52) has a unique, stable limit cycle surrounding the origin in the phase plane.

3) Poincare-Benedixson Theorem:

This theorem specifies conditions in 2D systems under which limit cycle exists. It is based on the idea that in a dynamic system $dx/dt = f(x)$, where $f(x)$ is continuous, no two trajectories intersect. Therefore, if there is a limit cycle, then all trajectories that start from inside the limit cycle will remain confined within the limit cycle. They can never come out of the limit cycle since to come out they have to cross the limit cycle, which is disallowed.

A formal statement of Poincare-Benedixson theorem is given below.

$\dot{x} = f(x)$ represents a continuously differentiable vector field on an open set containing a region R , where R is a closed, bounded subset of the plane, which is devoid of any fixed points, and

There exists a trajectory C that is “confined” in R (if it starts in R it stays in R forever).

If the above conditions are satisfied, C is a closed orbit.

It is not straightforward to apply Poincare-Benedixson theorem to a given system and show presence of limit cycles. Such demonstration depends crucially on construction of a trapping region, R , such that vector field on the borders of R points inwards everywhere.

There is a theorem from Index theory that any closed orbit must enclose one or more fixed points. Since Poincare-Benedixson requires that R must not include any fixed points, we

need to construct an R that excludes all fixed points. Typically a ring-like, annular region is chosen such that vector field is pointed inwards on the outer boundary, and outwards on the inner boundary. If such a region can be constructed, Poincare-Benedixson assures us that a limit cycle exists inside R . Exact construction of R depends on the details of the equations that describe the system.

2.3.5 Limit Cycle generation by Bifurcation:

As the parameters of a dynamic system are changed gradually, at certain critical values of the parameters, the dynamics can change qualitatively.(For example, in the map of fig. 10, a stable focus changes into an unstable focus when we cross the Δ –axis from below). Such qualitative changes in dynamics are known as bifurcations. The parameter values where such changes occurs are called bifurcation points.

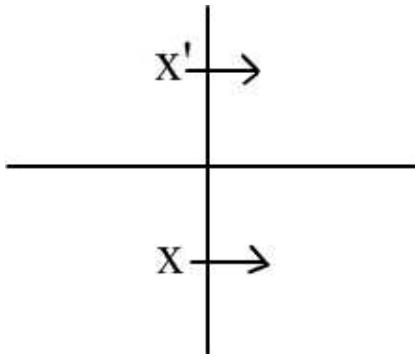
The map of Fig. 10 which shows different types of fixed points on the $\tau - \Delta$ plane is a simple example of bifurcations. The plane is divided into several regions, where each region represents one type of fixed point, or a dynamical regime. When we crossover from one region to a neighboring region, dynamics suddenly changes, and a bifurcation occurs. For example, when we cross the Δ –axis ($\tau = 0$ line), saddle node become either stable or unstable nodes.

Four important types of bifurcations by which limit cycles can be produced are:

- a) Supercritical Andronov-Hopf bifurcation
- b) Subcritical Andronov-Hopf bifurcation
- c) Saddle-node bifurcation
- d) Saddle-node on invariant circle bifurcation

a) **Andronov-Hopf Bifurcation – supercritical**

In Andronov-Hopf bifurcation, a stable focus gets converted into an unstable focus giving rise to a limit cycle in the process. A stable focus, as we know from the previous section, has complex valued eigenvalues with negative real parts; the eigenvalues of the unstable focus has positive real parts. Thus Andronov-Hopf bifurcation occurs when a pair of complex conjugate eigenvalues cross over from the left halfplane to the right.



Depending on the precise manner in which a stable focus gets destabilized and leads to a limit cycle, the Andronov-Hopf bifurcation can be classified into two types.

In supercritical Andronov-Hopf bifurcation, when a stable focus becomes a limit cycle, the limit cycle first starts with a small/infinitesimal amplitude and gradually increases in size.

This behavior can be seen in an example.

Example:

$$\dot{r} = \mu r - r^3 \quad (53)$$

$$\dot{\theta} = \omega + br^2 \quad (54)$$

Note that θ does not influence evolution of r (in Eqn. 53) but r influences θ (in eqn. 54).

Therefore changes in amplitude may be studied solely by studying Eqn. (53).

Note how the steady state value of r changes as μ is increased from a negative value to a positive value. For $\mu < 0$, the only real solution is $r = 0$; for $\mu > 0$, there are three real solutions: $r = 0$, $+\sqrt{\mu}$. Since r has to be positive, the solutions are $r = 0$ and $\sqrt{\mu}$. Eqn. (54) simply says that angular velocity increases with increasing radius.

Therefore, if we express the steady state radius, r_s , as a function of bifurcation parameter, μ ,

$$r_s = \begin{cases} \sqrt{\mu}, & \mu > 0 \\ 0, & \mu \leq 0. \end{cases} \quad \text{for } \mu > 0, \text{ and} \quad (55)$$

It can be easily verified that in the system defined by eqns. (53-54), the eigenvalues at the origin cross over from the left half-plane to the right, as μ crosses over from negative to positive values.

Eigen values analysis for eqns. (53-54)

$$\begin{aligned}
\dot{x} &= r \cos \theta - r \theta \sin \theta \\
&= (\mu v - r^3) \cos \theta - r(\omega + br^2) \sin \theta \\
&= (\mu - (x^2 + y^2))x - (\omega + b(x^2 + y^2))y \\
&= \mu x - \omega y + \text{cubic terms}
\end{aligned} \tag{56}$$

$$\begin{aligned}
y &= r \sin \theta \\
\dot{y} &= r \sin \theta + r \cos \theta \dot{\theta} \\
&= (\mu r - r^3) \sin \theta + r \cos \theta (\omega + br^2) \\
&= (\mu - r^2) y + x(\omega + br^2) \\
&= (\mu - x^2 - y^2)y + x(\omega + b(x^2 + y^2)) \\
&= \mu y + \omega x + \text{cubic terms}
\end{aligned} \tag{57}$$

$$A = \begin{bmatrix} \mu - \omega \\ \omega - \mu \end{bmatrix}$$

$$\lambda = \mu \pm i\omega \tag{58}$$

$\mu < 0$ – stable focus
 $\mu > 0$ – unstable focus

Properties of supercritical Andronov-Hopf bifurcation

- Size of the limit cycle, r_0 , grows as $\sqrt{\mu - \mu_c}$ for μ close to and greater than μ_c (Fig. 11)

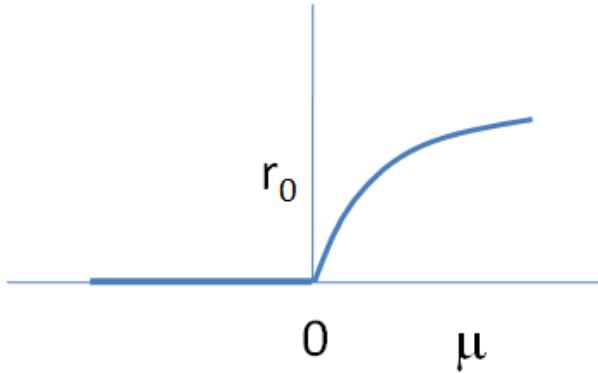
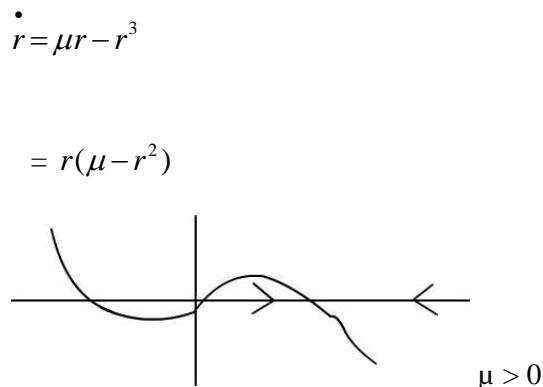


Figure 11: Size of the limit cycle r_o vs μ

2. Frequency $\approx \omega = I_m[\lambda]$, evaluated at $\mu = \mu_c$.
Formula is exact at birth of the limit cycle.

$$T = \frac{L\pi}{I_m[\lambda] + \theta(\mu - \mu_c)} \quad (59)$$



b) Subcritical Andronov-Hopf bifurcation:-

In a subcritical Andronov-Hopf bifurcation, when the unstable focus destabilized, a limit cycle of finite size suddenly appears, unlike in the supercritical case the limit cycle starts of small and grows gradually as the bifurcation parameter increases.

An example of a system which exhibits subcritical Andronov-Hopf bifurcation is given below:

$$\begin{aligned}\dot{r} &= \mu r + r^3 - r^5 \\ \dot{\theta} &= \omega + br^2\end{aligned}\tag{60}$$

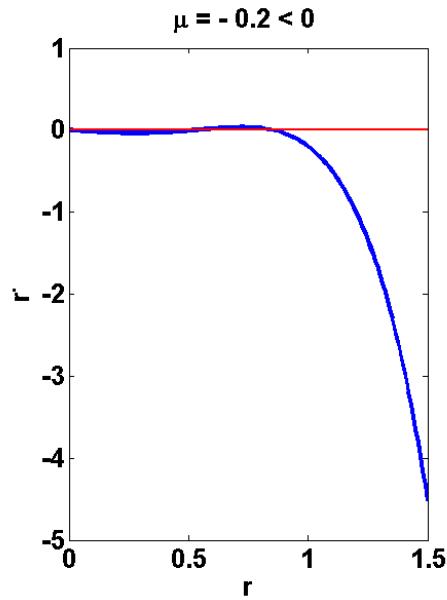


Figure 12: For $\mu < 0$

Thus for $\mu < 0$, there is no limit cycle; there is a stable fixed point at the origin (Fig. 12)

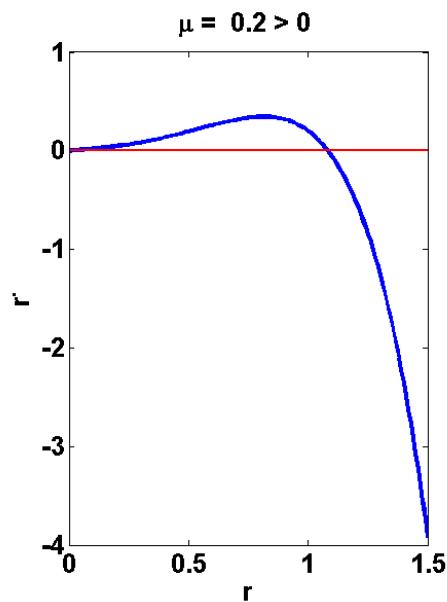


Figure 13: For $\mu > 0$

But when $\mu > 0$, there is suddenly a limit cycle at $r = r_0$, a finite value (Fig. 13)

Thus, in supercritical Andronov-Hopf bifurcation, the size of limit cycle varies as a function of μ as follows: (Fig. 14)

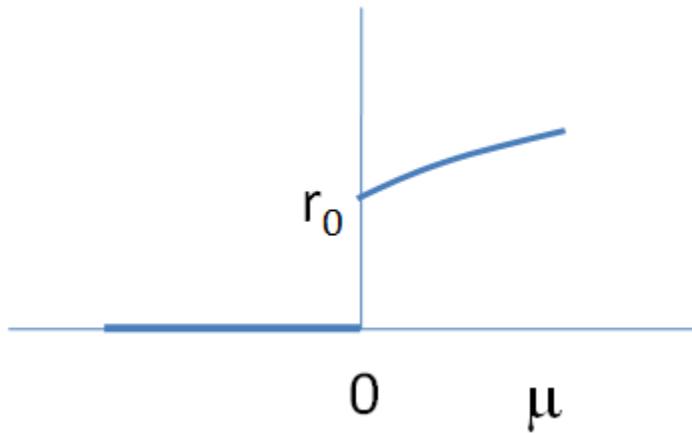


Figure 14: r_0 vs μ

c) Saddle node bifurcation of cycles:

In saddle node bifurcation of cycles (also called a fold bifurcation), two limit cycles coalesce and annihilate each other. A simple example system that exhibits Saddle node bifurcation of cycles is,

$$\dot{r} = \mu r + r^3 - r^5$$

$$\dot{\theta} = \omega + br^2$$

Note the system is the same as the one used in case of the subcritical Andronov-Hopf bifurcation above. But the difference lies in the value of bifurcation parameter at which the bifurcation occurs.

In case of subcritical Andronov-Hopf bifurcation, the bifurcation occurred at $\mu = 0$. For $\mu < 0$, there is a stable focus at the origin, an unstable limit cycle of lesser radius, and a stable limit cycle of larger radius. When $\mu > 0$, the unstable limit cycle merges with the origin and disappears. The stable limit cycle only remains.

In case of Saddle node bifurcation of cycles, the bifurcation occurs at $\mu = -1/4$. (Fig. 15)

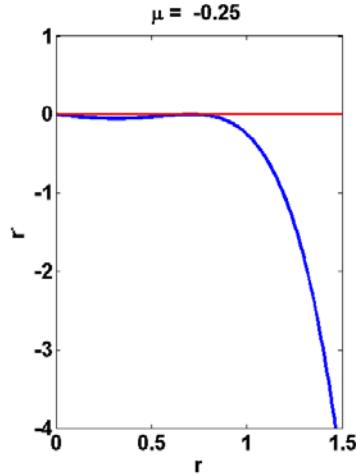


Figure 15: For $\mu = -0.25$

For $\mu < -1/4$ ($\mu < \mu_c$), (Fig. 16) (origin is the only stable point. There are no limit cycles (stable or unstable)).

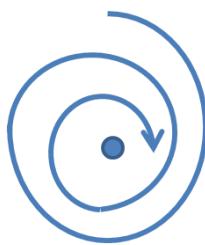


Figure 16: For $\mu < 0$

For μ ($\mu = \mu_c$) (Fig. 17),

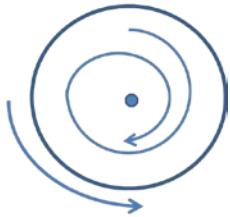


Figure 17: For $\mu = \mu_c$

For $(\mu > \mu_c)$: For $\mu > -1/4$, (Fig. 18) origin is still a stable point. But in addition there are now two limit cycles (one stable and the other unstable).

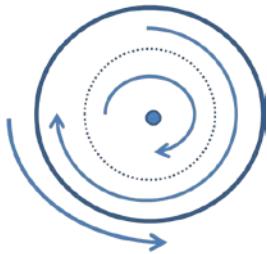


Figure 18: For $\mu > \mu_c$

Therefore a key difference between subcritical Andronov-Hopf bifurcation and Saddle node bifurcation of cycles is that in the former, the stable focus at the origin becomes unstable after bifurcation, whereas in the latter the focus at the origin remains stable throughout.

d) Saddle node on invariant Circle:

In this type of bifurcation, a node and saddle are located on a loop – the invariant circle. By varying a bifurcation parameter, the saddle and the node come together, coalesce and annihilate each other leaving the loop, the limit cycle, intact.

A sample system that displays Saddle node on invariant Circle behavior:

$$\begin{aligned}\dot{r} &= r(1-r^2) \\ \dot{\theta} &= \mu - \sin(\theta)\end{aligned}\tag{61}$$

For $\mu < 1$, the phase dynamic equation has two solutions – saddle and a node located on the unit circle ($r = 1$). Phase portrait is shown in the figure below. As $\mu \rightarrow 1$, the saddle and the node approach each other and coalesce. When $\mu > 1$, the saddle and the node annihilate each other leaving a limit cycle at $r = 1$. (Fig. 19)



Figure 19: Saddle node on invariant circle

Homoclinic Bifurcation:

A homoclinic orbit is one in which a trajectory starts at a saddle node, makes a loop, and returns to the same saddle node. In a homoclinic bifurcation, a closed loop approaches a saddle node. At bifurcation, the saddle touches the loop, transforming the loop into a homoclinic orbit. Thus trajectories that start from the saddle go along the loop and return to the saddle in the opposite direction (Fig. 20).

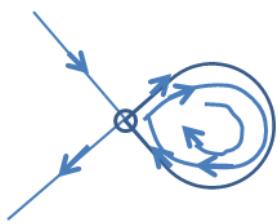


Figure 20: The limit cycle with the saddle outside

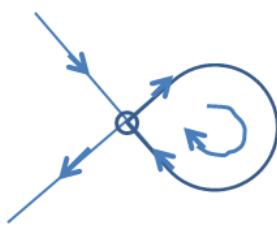


Figure 21: The saddle had merged with the limit cycle. The cycle has now become a homoclinic orbit.

Reference:

S. Strogatz., Nonlinear Dynamics and Chaos, Addison-Wesley publishing company, 1994.

3 Hodgkin-Huxley Model

The Hodgkin-Huxley model explains how the dynamics of ion channels (Na^+ , K^+ etc) contribute to the generation of an Action Potential in a neuron.

An Action Potential is a sharp voltage spike elicited by stimulating a neuron with a current that exceeds a certain threshold value. The current amplitude is increased gradually, at a threshold amplitude, the voltage response does not increase proportionally. It shows a sharp, disproportionate increase. Once the membrane voltage reaches a threshold value, it increases further rapidly to maximum value and drops again rapidly to a value that is less than resting value, before returning to the baseline value after a delay.

To describe the processes that lead to the generation of an AP from the introduction on Nernst Potential, we present a simple ion channel model which expresses how the channel conductance contributes to the membrane potential, and conversely, in case of voltage-sensitive channels, how the membrane potential controls channel conductance. Finally we place the above components in a full circuit model of neural membrane - the Hodgkin-Huxley model.

Now let us calculate the typical values of Nernst potentials of Na^+ and K^+ for a neuron.

3.1 Nernst potential:

Sodium Nernst potential:

Under resting conditions sodium concentration outside the neuron ($[\text{Na}^+]_o$) equals 440 mM, while inside ($[\text{Na}^+]_i$) it is 60 mM. From eqn. (1.1) we calculate Na^+ Nernst potential (E_{Na}) to be about 50 mV.

Potassium Nernst potential:

Similarly under resting conditions, $[\text{K}^+]_o$ equals 20 mM, while $[\text{K}^+]_i$ is 400 mM. Therefore, E_K is about -77 mV.

The Nernst potentials, like the membrane potential, are always measured inside, with the extracellular space as the reference. Also note that Na^+ Nernst potential is positive inside, and

K^+ Nernst potential is negative inside. In addition to Nernst potential, an ion channel also has a conductance, which is higher when it is in OPEN state than when it is closed. Channel conductance is usually expressed not in ‘per channel’ terms, but the total conductance of a whole patch of the membrane containing that channel, expressed in ‘Siemens/area’. Thus, again under resting conditions, Na^+ conductance is about 120 mS/cm^2 and K^+ conductance is about 36 mS/cm^2 .

3.2 Modeling the neural Membrane:

Electrical response of a neuron depends on the ion channels present and the membrane itself. Once we know the Nernst potential associated with a channel and the channel conductance we are in a position to build a basic model of an ion channel. The model is meant to capture the voltage-current (V-I) characteristics of the ion channel. Since a battery (Nernst potential) and a conductance are associated with the channel, we can represent the channel in of the following two ways: the battery in series (or parallel) (fig 3.2.1) with the conductance. Let us determine which of the two are correct. When the conductance is 0, as when the channel is closed, current through the channel is 0, irrespective of the value of the Nernst potential. This rule is satisfied only when the battery and conductance are in series and not in parallel. Therefore, as a first cut we represent an ion channel as a series of a battery and a conductance.

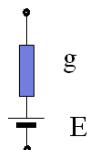


Figure 3.2.1: A basic model of an ion channel.

Since there are a variety of ion channels with distinct values of E and g , we have a separate branch for each of them. All these branches must be connected in parallel since all the channels stagger across the membrane, and share the same membrane voltage. In addition to the ion channels, the membrane itself is another electrical component that controls the dynamics of the membrane potential. The neural membrane is a lipid bilayer with insulating properties. Due to its bilayer structure, the membrane is modeled as a parallel plate capacitor. Thus the various ion channels and the membrane together may be depicted as an electrical equivalent circuit as follows:

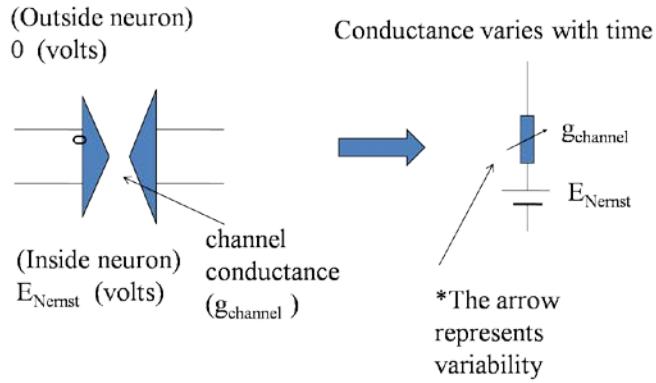


Figure 3.2.2: An ion channel has voltage dependant conductance.

Now if all the circuit components (capacitance, conductances, batteries) are time-invariant, then all the branches corresponding to ion channels can be combined into a single equivalent branch consisting of a single conductance and a battery. This can be done by invoking a well-known result from electrical engineering known as Thevenin's theorem. With such simplification a model of the membrane looks as shown in fig 3.2.3

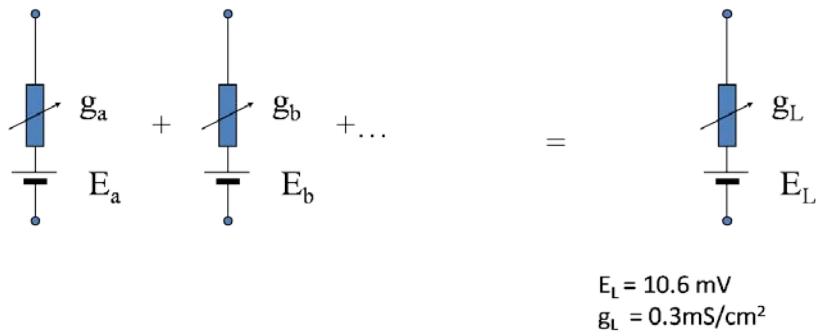


Figure 3.2.3: Resultant branch equivalent for time invariant branches, corresponding to ion channels

This is a simple, linear RC circuit and is not likely to exhibit the interesting voltage dynamics underlying AP generation, which is basically a nonlinear behavior. The root cause of AP behavior is presence of voltage-dependent ion channels, which may be regarded as nonlinear conductances since their conductance depends on membrane voltage. Particularly voltage-dependent Na⁺ and K⁺ channels play a crucial role in AP generation. These are shown in Fig. 3.2.4

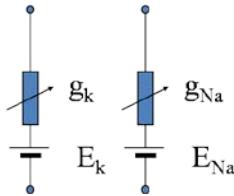


Figure 3.2.4: Circuit equivalents for sodium and potassium channels

We now take up the question of modeling a general voltage-sensitive channel first, followed by specific models for voltage-sensitive Na^+ and K^+ channels.

3.3 A general model of a voltage-sensitive channel:

Mathematical treatment of voltage dependent gating of channels has two parts. First, we describe the dynamics of switching between open/close states. This description is kept general, without reference to a specific gating mechanism. Next, we describe voltage-gating, or, the manner in which switching dynamics is influenced by membrane voltage.

3.3.1 Channel switching:

When we set out to describe switching between open and closed states of channel there is an implicit assumption that the channel has only two states. This assumption is not true. Complex channels do have a large number of intermediate states that are not strictly ‘open’ or ‘closed.’ Channel switching in those cases involves switching among all those states. But for the moment we consider only simple channels with only two states – open and closed.

Channel switching can described as a unimolecular chemical reaction where a molecule – the channel protein – switches between two states in fig 3.3.1.1

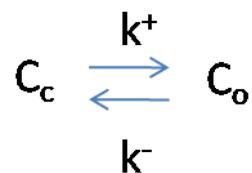


Figure 3.3.1.1: Channel Switching

where C_c and C_o denote the channel in open and closed states, and k^+ and k^- are the forward and reverse rate coefficients. If x is the fraction of the channels in open state, and $(1-x)$ is the fraction of channels in closed state, dynamics of state transition may be described as a first order chemical reaction as follows:

$$\frac{dx}{dt} = k^-x - k^+(1-x) \quad (3.3.1.1)$$

The above equation can also be written as,

$$\tau \frac{dx}{dt} = -x + x_\infty \quad (3.3.1.2)$$

where,

$$x_\infty = \frac{k^+}{k^+ + k^-}, \quad \tau_\infty = \frac{1}{k^+ + k^-} \quad (3.3.1.3)$$

Note that the above simple description is applicable for a statistical ensemble of channels. Since a cell membrane typically has a few thousand ion channels the above description is justified. Switching behavior in single channels, which can be measured using techniques like patch-clamping, is far more complex. Since our objective is to study membrane voltage changes it is sufficient to treat channels as an ensemble.

Complete solution of eqn. (3.3.1.1) requires knowledge of K_+ and K_- . In voltage-sensitive channels these quantities are functions of membrane voltage and that is the mechanism by which membrane voltage controls channel gating.

Since ion channels are proteins which contain charged amino acid side chains, membrane potential can influence the rate of open/close transitions. Using Arrhenius expressions, the rate constants can be expressed in terms of the membrane potential as:

$$k^+ = k_0^+ \exp(-\alpha V) \quad \text{and} \quad k^- = k_0^- \exp(-\beta V) \quad (3.3.1.4)$$

where k_0^+ and k_0^- are independent of membrane voltage.

Substituting the above formulae for k_0^+ and k_0^- in eqn. (3.3.1.2, 3.3.1.3) above, we have

$$x_\infty = \frac{1}{1 + (k_0^- / k_0^+) \exp((\alpha - \beta)V)} \quad (3.3.1.5)$$

$$\tau = \frac{1}{k_0^+ \exp(-\alpha V)} \frac{1}{1 + (k_0^- / k_0^+) \exp((\alpha - \beta)V)} \quad (3.3.1.6)$$

Now let us define,

$$S_0 = \frac{1}{(\beta - \alpha)} \quad \text{and} \quad V_0 = \frac{\ln(k_0^- / k_0^+)}{(\beta - \alpha)} \quad (3.3.1.7)$$

Substituting S_0 and V_0 in eqns. (3.3.1.5, 3.3.1.6) we obtain,

$$x_\infty = \frac{1}{1 + \exp(-(V - V_0) / S_0)} \quad (3.3.1.8)$$

$$\tau = \frac{\exp(\alpha V)}{k_0^+} \frac{1}{1 + \exp(-(V - V_0) / S_0)} \quad (3.3.1.9)$$

Using hyperbolic functions, the last two expressions can be rewritten as,

$$x_\infty = 0.5(1 + \tanh((V - V_0) / (2S_0))) \quad (3.3.1.10)$$

$$\tau = \frac{\exp(V(\alpha + \beta)/2)}{2\sqrt{k_0^+ k_0^-}(-(V - V_0)/S_0)} \quad (3.3.1.11)$$

Eqn. (3.3.1.10) tells us how many channels are open once the entire population of channels comes to equilibrium at a given steady state membrane voltage. Fig. 3.3.1.1 depicts the dependence of open probability (x) on membrane voltage for various values of V_0 and S_0 . If S_0 is positive, channels open with increasing membrane voltages. Such gates are known as *activation* gates. When S_0 is negative, gates close with increasing membrane voltages. Such gates are known as inactivation gates. V_0 determines the voltage at which the transition from ‘mostly open’ to ‘mostly closed’ takes place.

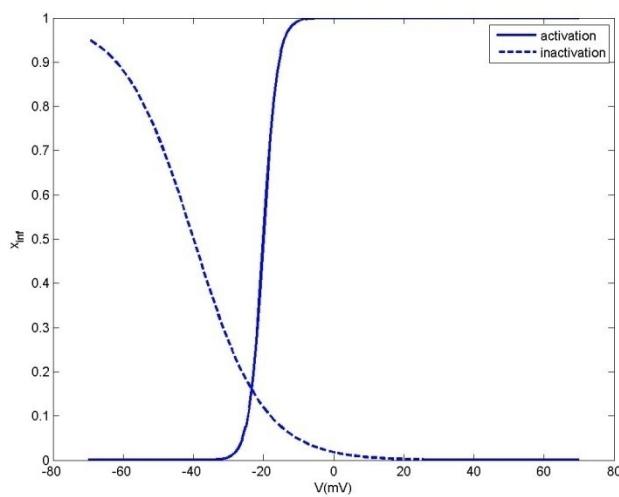


Fig. 3.3.1.1: the effect of membrane voltage (V) on x_∞ .

The solid line corresponds to an activation gate ($S_0 = 2$) and the dashed line corresponds to an inactivation gate ($S_0 = -10$). Note the smaller the magnitude of S_0 , the steeper the curve. V_0 for the activation and inactivation gates are -20 mV and -40 mV respectively.

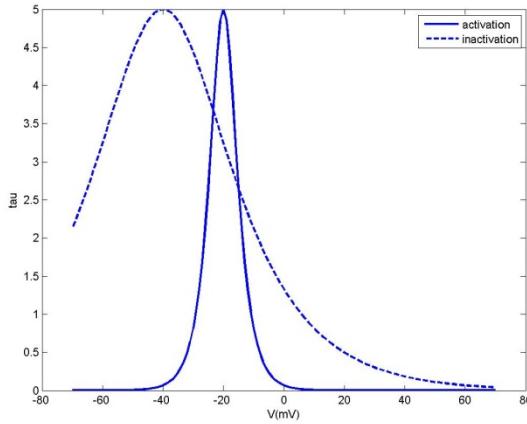


Fig.3.3.1.2 Effect of membrane voltage on the time constant of relaxation (t).

For the two curves shown in fig. 3.3.1.2, it is assumed that $\alpha + \beta = 0$, and

$$\phi = \frac{1}{2\sqrt{k_0^+ k_0^-}}, \text{ where } 2\sqrt{k_0^+ k_0^-} = 0.2 \text{ ms}^{-1}. \quad (3.3.1.12)$$

In fig. 3.3.1.2, for the solid line, $S_0 = 2$; $V_0 = -20$, whereas for the dotted line, $S_0 = 10$; $V_0 = -40$. Note that the sign of S_0 does not affect t since it is an even function of S_0 . V_0 denotes the voltage at which t is maximum. The magnitude of S_0 determines the sharpness of the peak.

3.3.2 Channel gating

In the above account of channel switching we have given the impression that there is a single “bottle-neck” in the pore which closes the channel. However, such a description is not completely accurate. There can be several local “bottle-necks,” simply referred to as *gates*, which can control channel closure.

To describe channel switching then we must be able to express the kinetics of all the gates in the channel. The question now is to describe the dynamics of single gates – the gate kinetics – and express the switching of the whole channel in terms of the component gates.

Our work is made easy by the fact that gate kinetics is also treated just as we have treated channel switching in eqns. (3.3.1.1 - 3.3.1.12). If we consider only channels with single gates, equations for gate kinetics would look identical to eqns. (3.3.1.1 - 3.3.1.12) with the only difference that ‘x’ now represents the fraction of open *gates*, which is also interpreted as *probability* of the gate being in open state. The same quantity is also called the *gating variable*. Since we are considering at the moment channels with only a single gate, fraction of open channels is the same as fraction of open gates.

Channels with multiple gates:

If a channel has K gates, with open probabilities of the gates denoted, by x_1, x_2, x_K , then the open probability, denoted by x, of the entire channel is given as:

$$x = x_1 * x_2 * \dots * x_K \quad \text{since the channel is open only when all the gates are open.}$$

3.3.3 Modeling a voltage-dependent Na+ channel:

The voltage-dependent Na+ channel in the Hodgkin-Huxley model is thought to have 4 gates – three of them being activation gates, and the last one being an inactivation gate. The three activation gates are thought to be identical, denoted by a common gating variable, m. The inactivation gate is denoted by the gating variable h. Thus the open probability of the entire gate is $m^3 h$. If the conductance of a population of Na+ channels in which all Na+ channels are fully open is g_{Na}^{max} then the conductance, g_{Na} of the population in general conditions (some Na+ channels are closed) is,

$$g_{Na} = g_{Na}^{max} m^3 h \quad (3.3.3.1)$$

The gate kinetics of m variables may be described as,

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m \quad (3.3.3.2)$$

Alternatively, the above equation may be written in the form of (3.3.1.2) in terms of the time constant and steady state value of the gating variable m . But the form of eqn. (3.3.3.2) above is more commonly used in literature. Similarly, the gate kinetics of h variable may be written as,

$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h \quad (3.3.3.3)$$

3.3.4 Modeling a voltage-dependent K+ channel:

The voltage-dependent K+ channel used in the HH model is thought to have 4 identical activation gates, denoted by the gating variable n . Thus the open probability of the entire gate is n^4 . If the conductance of a population of K+ channels in which all K+ channels are fully open is g_K^{max} then the conductance, g_K , of the population is,

$$g_K = g_K^{max} n^4 \quad (3.3.4.1)$$

Dynamics of the gating variable, n , is expressed as,

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n \quad (3.3.4.2)$$

The alpha and beta functions ($\alpha_m, \beta_m, \alpha_n, \beta_n, \alpha_h, \beta_h$) are estimated from voltage clamp experiments.

3.4 The Hodgkin-Huxley model

With this background, we are ready to introduce the Hodgkin-Huxley model equations. To this end, we redraw the circuit of fig. 3.2.3 with a slight modification. In the circuit of fig. 3.2.3, we have combined all the ion channels using the Thevenin's theorem. However, such compression is possible only when the all the components (conductances and batteries) are constant. But since conductances of voltage-sensitive channels vary through time, such compression is invalid. Thus we have to maintain the distinctness of voltage-sensitive Na+ and K+ channels. All the remaining channels, which are voltage-independent, can still be compressed to a single branch, consisting of a conductance and battery. This branch is thought to represent an notional ion

channel that is equivalent to the sum total of all the voltage-independent ion channels in the membrane. This ion channel is called a ‘leakage channel’ since it is constantly open, allowing ‘leakage’ of current from the neuron. We thus have four branches in the circuit: the capacitor, the voltage-sensitive Na⁺ channel, the voltage-sensitive K⁺ channel and the leakage channel (fig. 3.4.1). External current, I_{ext} , applied to the neuron can be split into four components:

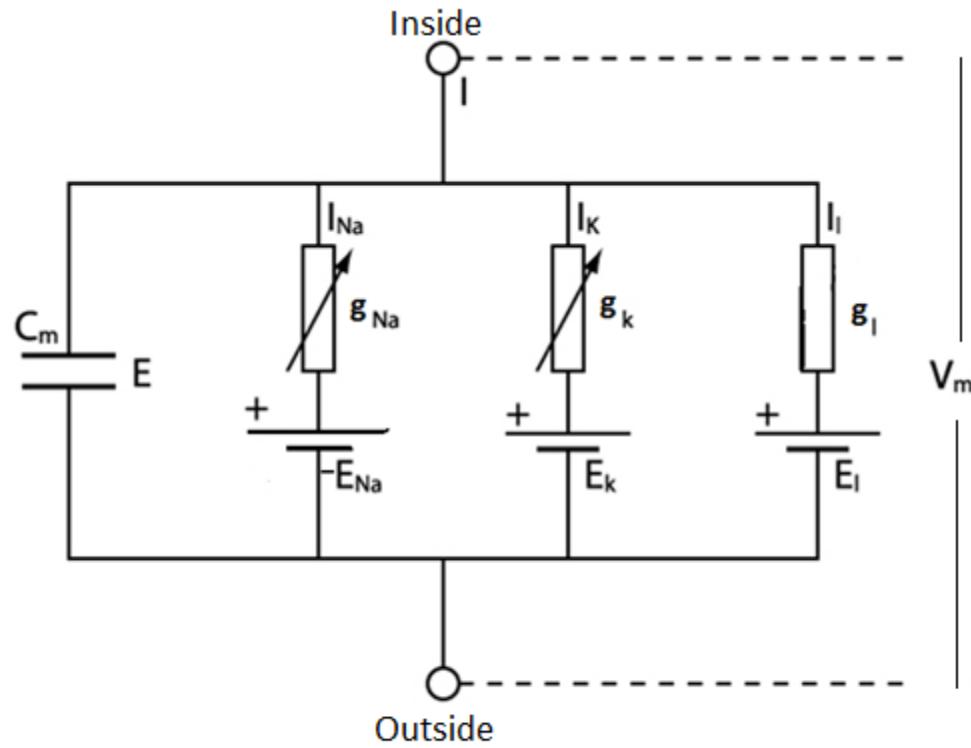


Figure 3.4.1: Equivalent circuit of a HH model

$$I_{ext} = I_C + I_{Na} + I_K + I_l \quad (3.4.1)$$

I_C – current through the capacitance

I_{Na} - current through the Na⁺ channel

I_K – current through the K⁺ channel

I_l – current through the leakage conductance

The above current equation may be expanded as,

$$C \frac{dV_m}{dt} + g_{Na}(V_m - E_{Na}) + g_K(V_m - E_K) + g_l(V_m - E_l) = I_{ext}$$

Or,

$$C \frac{dV_m}{dt} = -g_{Na}^{\max} m^3 h (V_m - E_{Na}) - g_K^{\max} n^4 (V_m - E_K) - g_l (V_m - E_l) + I_{ext} \quad (3.4.2)$$

The above equation that describes membrane voltage dynamics must be supplemented by the following equations to complete the definition of the Hodgkin-Huxley model.

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m \quad (3.4.3)$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h \quad (3.4.4)$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n \quad (3.4.5)$$

$$g_{Na} = g_{Na}^{\max} m^3 h \quad (3.4.6)$$

$$g_K = g_K^{\max} n^4 \quad (3.4.7)$$

$$\alpha_n = \frac{0.01(v_m + 50)}{\left\{1 - \exp\left(-\frac{[v_m + 50]}{10}\right)\right\}}, \quad \beta_n = 0.125 \exp\left(-\frac{[v_m + 60]}{80}\right) \quad (3.4.8)$$

$$\alpha_m = \frac{0.1(v_m + 35.0)}{\left\{1 - \exp\left(-\frac{[v_m + 35]}{10}\right)\right\}}, \quad \beta_m = 4 \exp\left(-0.0556[v_m + 60]\right) \quad (3.4.9)$$

$$\alpha_m = 0.07 \exp(-0.05[v_m + 60]), \quad \beta_h = \frac{1}{(1+\exp\{-0.1[v_m + 30]\})} \quad (3.4.10)$$

Dependencies among the variables involved in the HH model are depicted in fig. 3.4.2

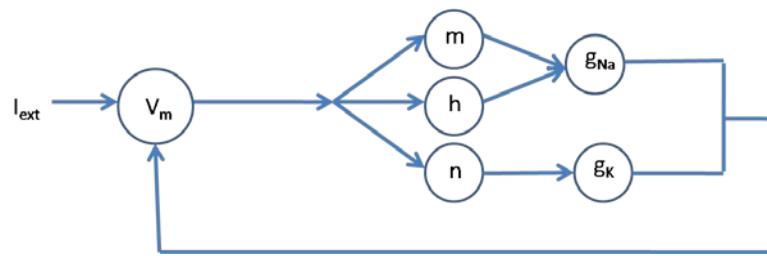


Figure 3.4.2: Dependencies among the variables involved in the HH model

The Voltage, Conductances, gating variables vs. time for different I values showing various dynamics have been simulated below (figs 3.4.3 - 3.4.6). Fig 3.4.7 shows the complete frequency dynamics of the HH model.

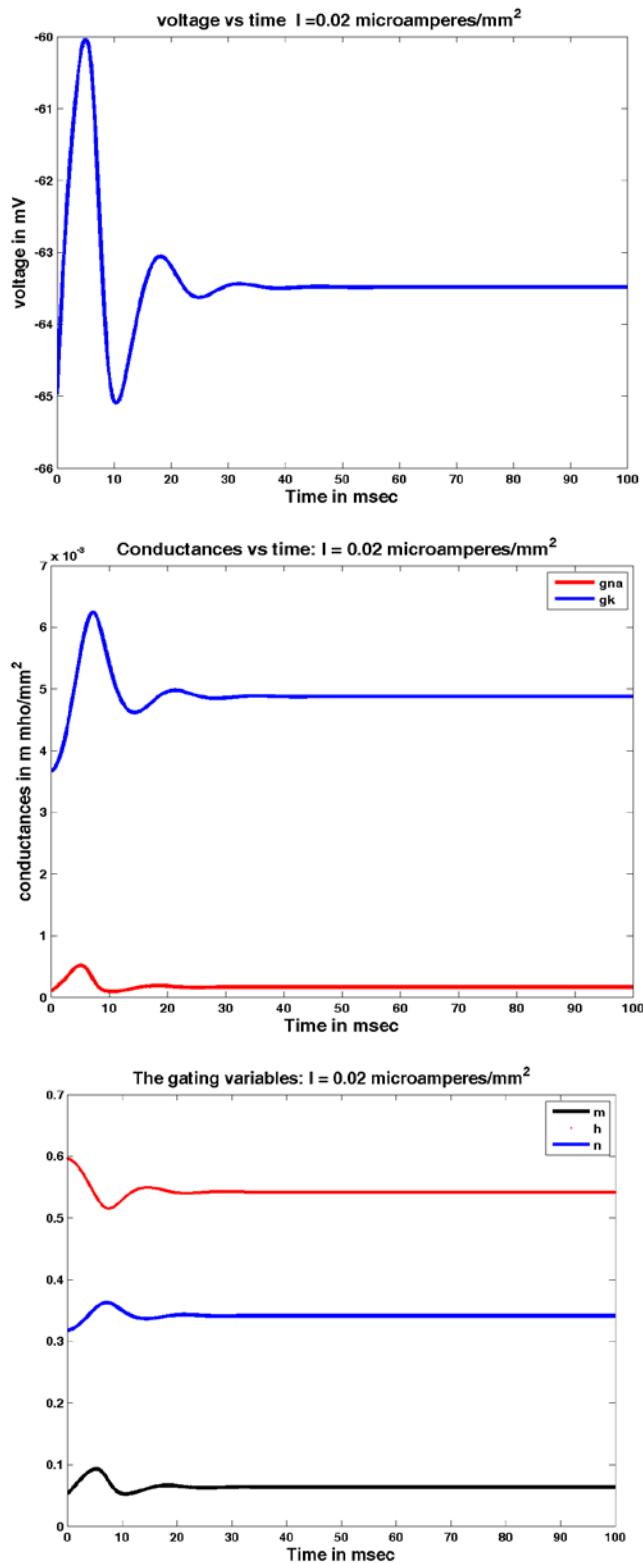


Figure 3.4.3: Voltage vs. time; Conductances vs.time;gating variables vs. time for $I = 0.02$ $\mu A/mm^2$

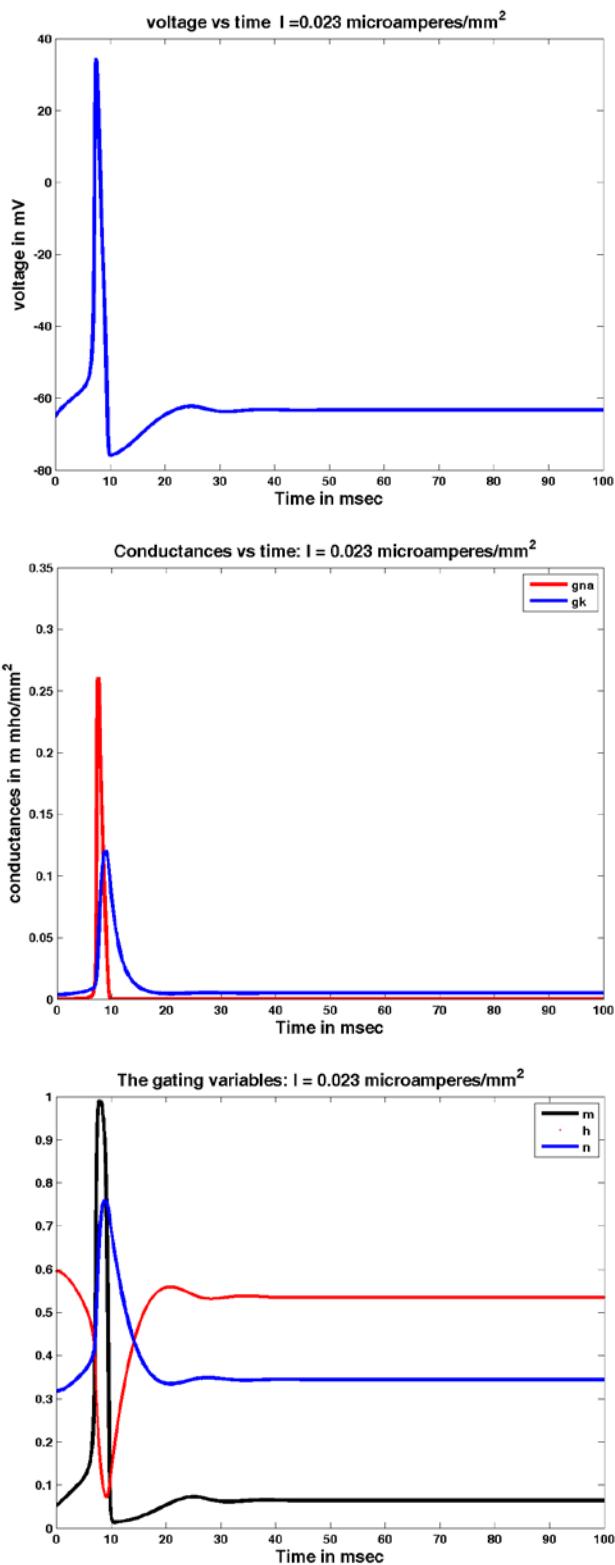


Figure 3.4.4: Voltage vs. time; Conductances vs. time;gating variables vs. time for $I = 0.023 \mu\text{A/mm}^2$

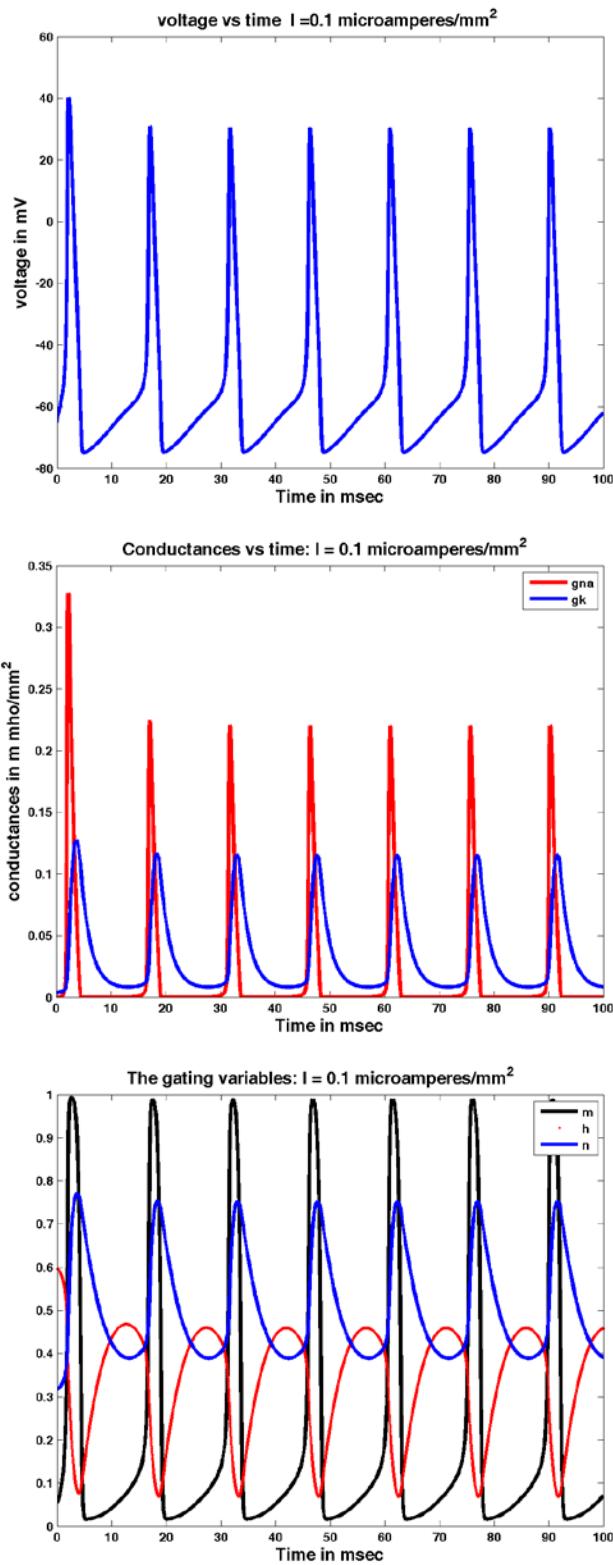


Figure 3.4.5: Voltage vs. time; Conductances vs. time;gating variables vs. time for $I = 0.1$

$$\mu\text{A/mm}^2$$

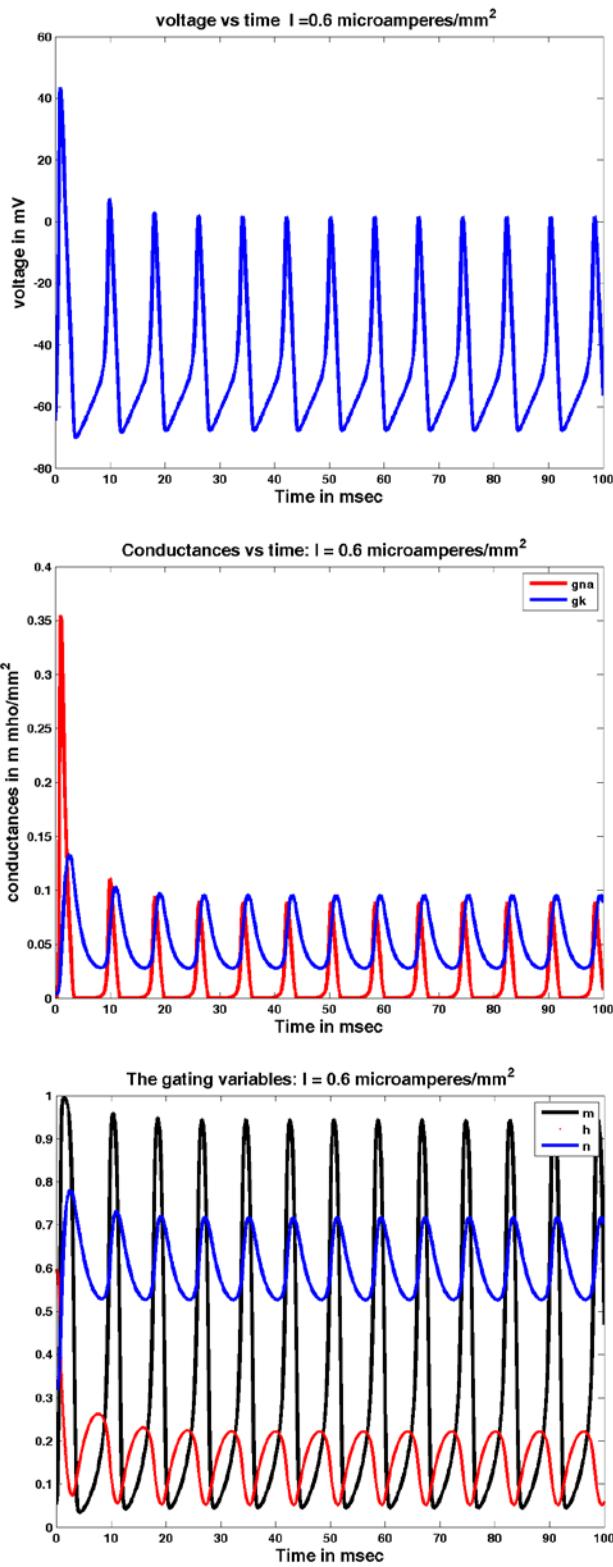


Figure 3.4.6: Voltage vs. time; Conductances vs. time;gating variables vs. time for $I = 0.6$

$$\mu\text{A}/\text{mm}^2$$

The HH model Action potential dynamics see the following fixed points. Before I1, no AP's are seen. Between I1 and I2, finite number of Action potentials are seen, and between I2 and I3, Limit cycle behavior is noticed. After I3, no Action potentials are seen.

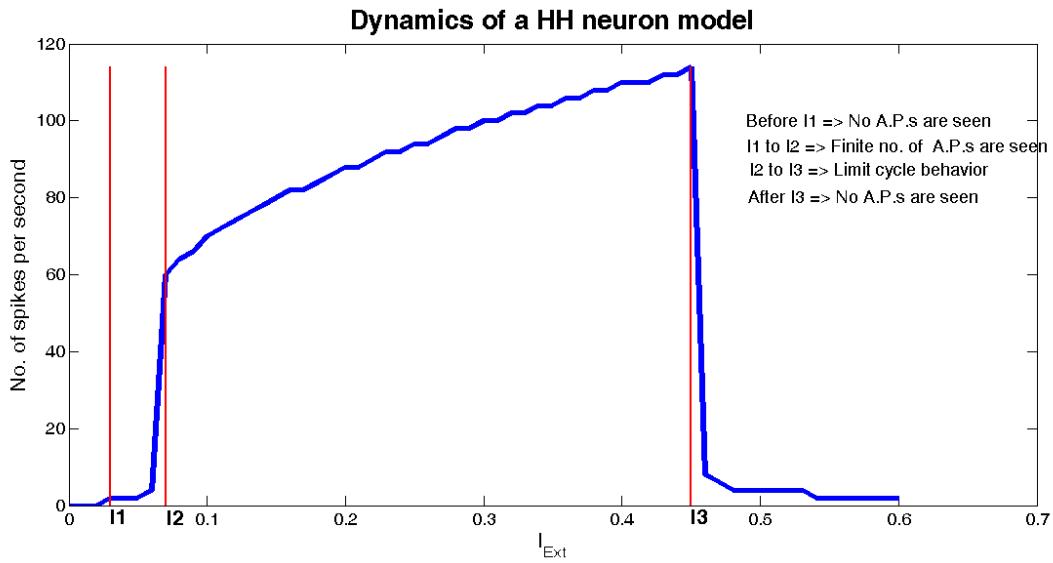


Figure 3.4.7: Dynamics of HH neuron model

Code:

```
%THIS PROGRAM DEMONSTRATES HODGKIN HUXLEY MODEL IN CURRENT CLAMP EXPERIMENTS
AND SHOWS ACTION POTENTIAL PROPAGATION
%Time is in msec, voltage in mvs, conductances in m mho/mm^2, capacitance in
microF/mm^2
% threshold value of current is 0.0223

k=1;
istep=0.01;
for ImpCur=0:istep:0.6
%TimeTot=input('enter the time for which stimulus is applied in
milliseconds');

gkmax=.36;
vk=-77;
gnamax=1.20;
vna=50;
gl=0.003;
vl=-54.387;
cm=.01;
```

```

dt=0.01;
niter=50000;
t=(1:niter)*dt;
iapp=ImpCur*ones(1,niter);
%for i=1:100
%    iapp(1,i)=ImpCur;
%end;
v=-64.9964;
m=0.0530;
h=0.5960;
n=0.3177;

gnahist=zeros(1,niter);
gkhist=zeros(1,niter);
vhist=zeros(1,niter);
mhist=zeros(1,niter);
hhist=zeros(1,niter);
nhist=zeros(1,niter);

for iter=1:niter
gna=gnamax*m^3*h;
gk=gkmax*n^4;
gtot=gna+gk+gl;
vinf = ((gna*vna+gk*vk+gl*vl)+ iapp(iter))/gtot;
tauv = cm/gtot;
v=vinf+(v-vinf)*exp(-dt/tauv);
alpham = 0.1*(v+40)/(1-exp(-(v+40)/10));
betam = 4*exp(-0.0556*(v+65));
alphan = 0.01*(v+55)/(1-exp(-(v+55)/10));
betan = 0.125*exp(-(v+65)/80);
alphah = 0.07*exp(-0.05*(v+65));
betah = 1/(1+exp(-0.1*(v+35)));
taum = 1/(alpham+betam);
tauh = 1/(alphah+betah);
taun = 1/(alphan+betan);
minf = alpham*taum;
hinf = alphah*tauh;
ninf = alphan*taun;
m=minf+(m-minf)*exp(-dt/taum);
h=hinf+(h-hinf)*exp(-dt/tauh);
n=ninf+(n-ninf)*exp(-dt/taun);
vhist(iter)=v; mhist(iter)=m; hhist(iter)=h; nhist(iter)=n;
gnahist(iter) = gna;
gkhist(iter) = gk;
end

j=1;
realpeaks=zeros;
[peaks, locs]=findpeaks(vhist);
for temp=1:length(peaks)
if peaks(temp) >=10 % minimum value at which a waveform is considered AP.
realpeaks(j)=peaks(temp);
j=j+1;
end;

```

```

end;
if realpeaks ~= 0
no_of_peaks(k)=length(realpeaks);
else
no_of_peaks(k)=0;
end;
k=k+1
end;
figure(1)
%subplot(2,1,1)
plot(t,vhist)
title('voltage vs time')

figure(2)
%subplot(2,1,2)
plot(t,mhist,'y-', t,hhist,'g.',t,nhist,'b-')
legend('m','h','n')

figure(3)
gna=gnamax*(mhist.^3).*hhist;
gk=gkmax*nhist.^4;
clf
plot(t,gna,'r');
holdon
plot(t,gk,'b');
legend('gna','gk')
holdoff

figure(4);
X=0:istep:0.6;
plot(X,no_of_peaks*1000/(niter/100));
xlabel('I_{Ext}');
ylabel('No. of spikes per second')
holdon;
for l=2:length(no_of_peaks) %to define I1, I2, I3.
if no_of_peaks(l)>0 && no_of_peaks(l-1)==0
    I1=(l-1)*istep
end;
if no_of_peaks(l)>no_of_peaks(l-1)+4
    I2=(l-1)*istep
end;
if no_of_peaks(l)<no_of_peaks(l-1)-2
    I3=(l-2)*istep
end;
end;
I1=0*no_of_peaks*1000/(niter/100)+I1;
plot(I1,no_of_peaks*1000/(niter/100),'r');
text(I1(1),-3,'I1');
I2=0*no_of_peaks*1000/(niter/100)+I2;
plot(I2,no_of_peaks*1000/(niter/100),'g');
text(I2(1),-3,'I2');
I3=0*no_of_peaks*1000/(niter/100)+I3;
plot(I3,no_of_peaks*1000/(niter/100),'y');
text(I3(1),-3,'I3');
text(0.5,100,'Before I1 => No A.P.s are seen');
text(0.5,95,'I1 to I2 => Finite no. of A.P.s are seen');

```

```
text(0.5,90,'I2 to I3 => Infinite no.of A.P.s are seen');
text(0.5,85,'After I3 => No A.P.s are seen');
```

4 Modeling the neuron components

The aim of this chapter is to present the modeling components that would constitute a model of a whole neuron. We have described the anatomy of a neuron in chapter 2. We have discussed the electrophysiological basis of electrical signaling that occurs in a neuron in chapter 3. In this chapter we combine the biological background presented in the two fore-mentioned chapters and present modeling equations.

To begin with let us quickly recall the four components of neural signaling or, rather, the four stages of a neural signal in its passage from the “input” (apical dendrite) of a neuron to its “output” (axon terminal).

- 1) signal propagation along the dendrite towards the soma
- 2) spatial and temporal summation in the soma
- 3) signal propagation along the axon
- 4) neurotransmission across the synapse

We had also noted earlier that signal propagation along dendrites is mostly passive, as along an electrical cable; that summation occurs in the axon hillock; that an intact action potential propagates down the axon without losing amplitude because it is charged all along the way by voltage-sensitive channels; that neurotransmission occurs across a synapse – as though there is a “hotline” from axon terminal A to apical dendrite B – via chemical means; and, finally, that this whole sequence of events occurs in a neat unidirectional fashion from the apical dendrites to axon terminals.

4.1 Dendrite

We introduce three electrical parameters of a dendritic cable. The parameters are defined per unit length of the cable.

1. Axial resistance: Resistance offered by the intracellular compartment per unit length of the cable of diameter, d . The resistivity of the intracellular medium is R_i .

We now relate the resistivity, R_i , which is an intrinsic property of the intracellular medium, to axial resistance, r_a , which is resistance per unit length of the cable.

In general the resistance, R , and resistivity, ρ , of a pipe of area of cross-section, A , and length, L , are related as:

$$R = \rho L/A, \text{ and}$$

Resistance per unit length of the pipe is:

$$= \rho/A$$

A similar relation for our cable is:

$$r_a = \frac{R_i}{A} = \frac{R_i}{\pi d^2 / 4} = \frac{4R_i}{\pi d^2} \quad (\Omega/\text{cm}) \quad (4.1.1)$$

2. Membrane resistance: The membrane offers resistance for flow current between the intracellular compartment and the extracellular space. This resistance is inversely proportional to the surface area of the membrane.

Therefore, if R_m is the resistance of a membrane patch of unit area, a quantity referred to as specific resistance, the total resistance offered by a cylinder of diameter, d , and unit length, is given as:

$$r_m = \frac{R_m}{A} = \frac{R_m}{\pi d \cdot 1} = \frac{R_m}{\pi d} \quad (\Omega\text{-cm}) \quad (4.1.2)$$

3. Membrane capacitance: The plasma membrane has a specific capacitance, C_m , of about 10^{-6} F/cm^2 . Therefore, capacitance of the cable of unit length, c_m , is,

$$c_m = C_m \pi d \quad (\text{F/cm}) \quad (4.1.3)$$

Using the electrical parameters defined above, we can now represent the cable as an electric circuit. In this circuit, the continuous cable is represented as a series of discrete circuit elements, in which each element approximates a short length of the cable, say, of length, Δx .

4.1.1 Infinite Cable

Applying Ohm's law to one of the horizontal resistances, $r_a \Delta x$,

$$V_m(x, t) - V_m(x + \Delta x, t) = \Delta x r_a I_i(x, t)$$

$$-\frac{\partial V_m}{\partial x} = r_a I_i(x, t) \quad (4.1.1.1)$$

Applying the law of continuity of current at a given node in the circuit of fig. 4.1.1,

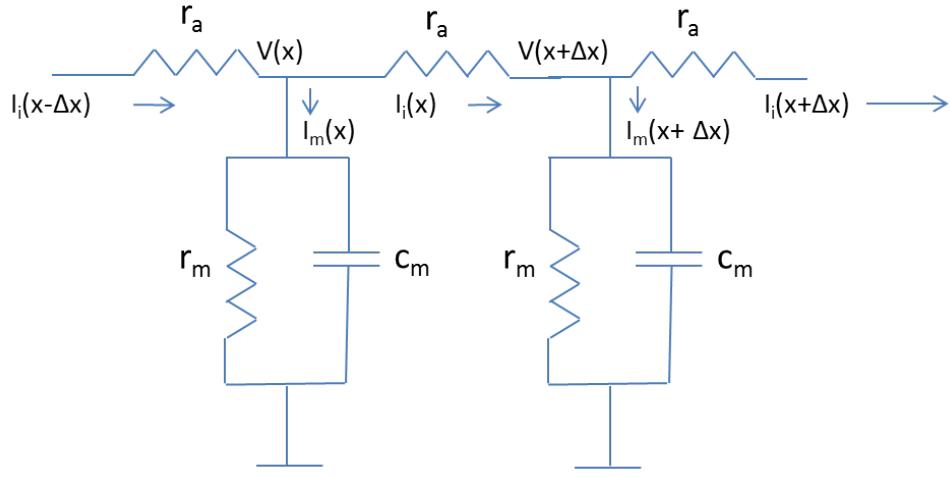


Figure 4.1.1: Circuit equivalent for a dendrite

$$I_i(x, t) - I_i(x - \Delta x, t) = -\Delta x I_m(x, t)$$

$$-\frac{\partial I_i}{\partial x} = I_m(x, t) \quad (4.1.1.2)$$

Combining (4.1.1.1) and (4.1.1.2),

$$\frac{1}{r_a} \frac{\partial^2 V_m}{\partial x^2} = I_m(x, t) \quad (4.1.1.3)$$

Now, the membrane current, I_m , can be resolved into three components 1) current through membrane capacitance, 2) current through membrane resistance, and 3) externally injected current, I_{ext} , if any. Thus,

$$I_m(x, t) = \frac{V_m - V_{rest}}{r_m} + c_m \frac{\partial V_m}{\partial t} - I_{ext} \quad (4.1.1.4)$$

Combining (4.1.1.3) and (4.1.1.4),

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} = V_m - V_{rest} + \tau_m \frac{\partial V_m}{\partial t} - r_m I_{ext} \quad (4.1.1.5)$$

where

$$\lambda = \sqrt{\frac{r_m}{r_a}} \quad \text{known as the space constant, and,}$$

$$\tau_m = r_m c_m \quad \text{known as the time constant, of the cable.}$$

Equation 4.1.1.5 is known as the Linear Cable Equation. Eqn. 4.1.1.5 can be further simplified if the membrane voltage, V_m , is defined with reference to the resting potential. Assuming that the resting potential is the same everywhere along the cable, it only offsets the membrane potential and does not affect the derivative terms in eqn. (4.1.1.5). Thus, from now on, if we designate V_m to represent the deviation of membrane potential from the resting potential, V_{rest} , the $(V_m - V_{rest})$ term in eqn. (4.1.1.5) can be replaced by, V_m , and we have the following simpler form.

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} = V_m + \tau_m \frac{\partial V_m}{\partial t} - r_m I_{ext} \quad (4.1.1.6)$$

Steady State Analysis:

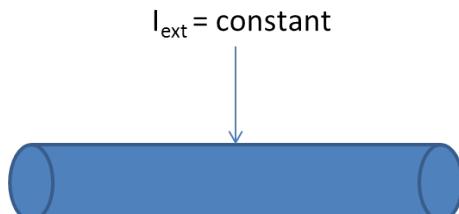


Figure 4.1.2: Infinite cable

Though our ultimate objective is to be able to describe signal transmission along the cables with complex geometries, we begin with a simple situation.

We consider an infinite cable in which a constant current is injected at a point. The goal is to determine membrane voltage distribution under steady state conditions.

External Current:

$$I_{ext} = I_0 \delta(x) u(t), \quad (4.1.1.7)$$

spatially it is a point source; and temporally it is a step function.

Boundary conditions:

$$V(x,t) = 0, \text{ at } |x| \rightarrow \infty, \forall t \quad (4.1.1.8.a)$$

Initial condition:

$$V(x,0) = 0, \quad x \in (-\infty, \infty) \quad (4.1.1.8.b)$$

Under steady state conditions,

$$\frac{\partial V_m}{\partial t} = 0$$

Therefore, eqn. (4.1.1.6) becomes,

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} = V_m - r_m I_{ext} \quad (4.1.1.9)$$

Solution to eqn. (4.1.1.9) will be of the form:

$$V_m(x) = A e^{x/\lambda} + B e^{-x/\lambda} \quad (4.1.1.10)$$

Since we are concerned with only steady state behavior, time is omitted, and membrane voltage is represented as, $V_m(x)$.

Now let us apply the boundary conditions eqn. (4.1.1.8), to the solution eqn. (4.1.1.10). Since $V_m(x)$ tends to 0 at $+\infty$, $A=0$, and since $V_m(x)=0$ at $-\infty$, $B = 0$. This difficulty can be overcome if we let the form of the solution be,

$$V_m(x) = V_0 e^{-|x|/\lambda} \quad (4.1.1.11)$$

where V_0 is the steady state voltage at $x=0$.

Let us try to verify that $V_m(x)$ of eqn. (4.1.1.11) satisfies eqn. (4.1.1.9).

$$\frac{\partial V_m(x)}{\partial x} = -\frac{V_0}{\lambda} e^{-|x|/\lambda} sign(x)$$

$$\frac{\partial^2 V_m(x)}{\partial x^2} = \frac{V_0}{\lambda^2} e^{-|x|/\lambda} - 2 \frac{V_0}{\lambda} e^{-|x|/\lambda} \delta(x) = \frac{V_0}{\lambda^2} e^{-|x|/\lambda} - 2 \frac{V_0}{\lambda} \delta(x)$$

$$\lambda^2 \frac{\partial^2 V_m(x)}{\partial x^2} = V_0 e^{-|x|/\lambda} - 2V_0 \lambda \delta(x)$$

Comparing the last equation with eqn. (4.1.1.9), we have

$$2\lambda V_0 = I_0 r_m$$

Or,

$$V_0 = \frac{I_0 r_m}{2\lambda} \quad (4.1.1.12)$$

Therefore, the final form of steady state membrane voltage of an infinite cylinder is,

$$V_m(x) = \frac{I_0 r_m}{2\lambda} e^{-|x|/\lambda} \quad (4.1.1.13)$$

where,

$$\lambda = \sqrt{\frac{r_m}{r_a}} = \sqrt{\frac{R_m}{R_i} \cdot \frac{d}{4}}$$

Electrotonic Distance:

Any length, l , can be expressed as electrotonic distance, L , as,

$$L = \frac{l}{\lambda} \quad (4.1.1.14)$$

Input resistance:

Input resistance, R_{in} , is defined as the ratio of voltage at the point of current injection to the magnitude of current injected.

$$R_{in} = \frac{V(x=0)}{I_0} = \frac{r_m}{2\lambda} = \frac{\sqrt{r_a r_m}}{2} \quad (4.1.1.15)$$

Example:

Consider a infinite cable with the static current $I_0 = 10$ mA injected at point $x=2$. Plot the static solution for such a cable.

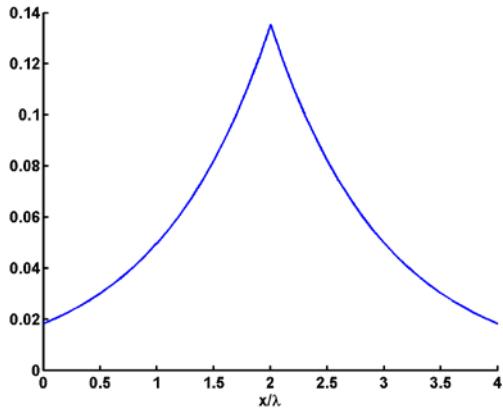
Solution:

The static solution for the infinite cable is given by: $V_m(x) = \frac{I_0 r_m}{2\lambda} e^{-|x|/\lambda}$.

Since current $I_0 = 10$ mA is injected at $x=2$,

$$I_{ext} = I_0 \delta(x);$$

The plot of $2V_m(x)/I_0 r_m$ vs x/λ would be:



4.1.2 Semi-infinite Cable:

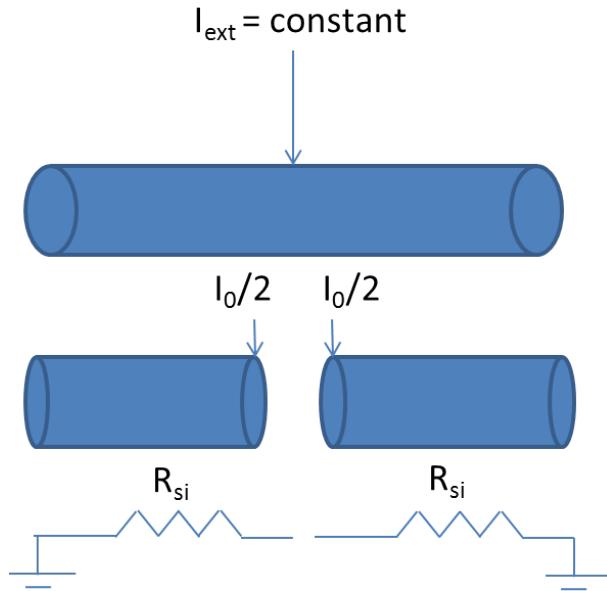


Figure 4.1.2.1: Semi-infinite cable

Let us consider the case of a semi-infinite cable where current of magnitude, I_0 , is injected into one end of the cylinder. Since an infinite cable can be viewed as two semi-infinite cables in parallel, steady state membrane voltage of a semi-infinite cable is twice that of the infinite cable, and is given as,

$$V_m(x) = \frac{I_0 r_m}{\lambda} e^{-|x|/\lambda} \quad (4.1.2.1)$$

Similarly, the input resistance, which is naturally twice that of an infinite cable, is,

$$R_{in} = \frac{V(x=0)}{I_0} = \frac{r_m}{\lambda} = \sqrt{r_a r_m} = r_a \lambda \equiv R_\infty \quad (4.1.2.2)$$

R_∞ is called the input resistance of a semi-infinite cable.

4.1.3 Finite Cable:

Consider a finite cable of electrotonic length L . Let X denote the electrotonic distance along the cable,

$$X = \frac{x}{\lambda} \quad (4.1.3.1)$$

A general expression for steady state membrane voltage, as a function of X , is given as,

$$V_m(X) = A \cosh(L - X) + B \sinh(L - X) \quad (4.1.3.2)$$

where,

$$\cosh(x) = (e^x + e^{-x})/2$$

$$\sinh(x) = (e^x - e^{-x})/2$$

Now we consider three different boundary conditions under which we solve the steady state voltage distribution of a finite length cable.

4.1.3.1 Sealed-end Boundary Condition:

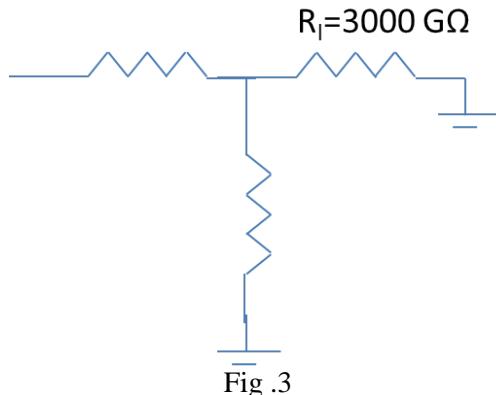


Figure 4.1.3.1: Sealed end boundary condition

Physically this refers to the situation when the dendrite is sealed/closed with a membrane patch. Electrically this is equivalent to loading the circuit of fig. 4.1.1 at the far end with a resistance equal to that of membrane patch that sealed the dendritic cable. For dendritic cable of diameter, d of $2\mu\text{m}$, and R_m of $10^5 \Omega\text{cm}^2$, the loading resistance R_L is,

$$R_L = 3000 \text{ G}\Omega$$

$$R_L = \frac{R_m}{\pi d^2 / 4} = 3000 G\Omega$$

Since the sealed end has high resistance we can approximate it with an open circuit, implying that axial current, $I_i = 0$ at the sealed end ($X=L$).

Therefore, from eqn. (4.1.1.1),

$$\left. \frac{\partial V_m}{\partial X} \right|_{X=L} = 0$$

Substituting the above boundary condition in the solution of eqn. (4.1.3.2), we have,

$$\left. \frac{\partial V_m}{\partial X} \right|_{X=L} = -A \sinh(L-X) - B \cosh(L-X) = 0$$

or,
 $B=0$.

Therefore,

$$V_m(X) = A \cosh(L-X)$$

Assuming, V_0 , is the voltage at the near-end ($x=0=X$) of the cable at steady state, the solution can be written as,

$$V_m(X) = V_0 \cosh(L-X) / \cosh(L) \quad (4.1.3.1.1)$$

$$I_i = -\frac{1}{r_a} \frac{\partial V_m}{\partial x} = \frac{V_0}{\cosh(L)} (-\sinh(L-X)) \left(-\frac{1}{r_a \lambda} \right) = \frac{V_0}{R_\infty} \frac{\sinh(L-X)}{\cosh(L)}$$

At $X=0$, I_i is,

$$= \frac{V_0}{R_\infty} \tanh(L) \quad (4.1.3.1.2)$$

$$\therefore R_{in} = R_\infty \coth(L) \quad (4.1.3.1.3)$$

4.1.3.2 Killed End Boundary Condition

The previous case of ‘sealed-end boundary condition’ refers to the situation when the far end is an open circuit. The present case of ‘killed end boundary condition’ the end of the terminal is ‘shorted’ so that the voltage at the far end is zero. Physically this situation can be created by cutting (“killing”) the far end so that the interior of the dendrite is directly in contact with the extracellular space.

The form of the solution for this case is again,

$$V(X) = A \cosh(L - X) + B \sinh(L - X)$$

Boundary conditions:

$$V(X) = 0 \text{ at } X = L$$

$$\therefore A = 0$$

$$V(X) = V_0 \text{ at } X = 0$$

$$\therefore B = \frac{V_0}{\sinh(L)}$$

Input resistance: $V(X)/I_i(X)$ at $X = 0$.

Input current is,

$$I_i = -\frac{1}{r_a} \frac{\partial V}{\partial x} \Big|_{X=0}$$

$$\frac{\partial V}{\partial X} = -V_0 \frac{\cosh(L - X)}{\sinh(L)}$$

$$\therefore I_i = -\frac{1}{r_a} \frac{\partial V}{\partial x} \Big|_{x=0} = \frac{V_0}{r_a} \coth(L)$$

$$\therefore R_m = \frac{V_0}{I_i} = \lambda r_a \tanh(L) = R_\infty \tanh(L)$$

Arbitrary Boundary Condition:

Let load resistance be $= R_L$.

Load voltage $= V_L$.

$$V(X) = A \cosh(L - X) + B \sinh(L - X)$$

Boundary conditions:

$$V(X) = V_L \text{ at } X = L$$

$$\therefore A = V_L$$

$$V(X) = V_0 \text{ at } X = 0$$

$$\therefore B = \frac{V_0 - V_L \cosh(L)}{\sinh(L)}$$

Substituting...

$$\begin{aligned}
 V(X) &= V_L \cosh(L-X) + \frac{(V_0 - V_L \cosh(L))}{\sinh(L)} \sinh(L-X) \\
 &= \frac{V_L \cosh(L-X) \sinh(L) + (V_0 - V_L \cosh(L)) \sinh(L-X)}{\sinh(L)} \\
 &= \frac{V_L \cosh(L-X) \sinh(L) + V_0 \sinh(L-X) - V_L \cosh(L) \sinh(L-X)}{\sinh(L)} \\
 (\text{Using, } \sinh(X) &= \sinh(L - (L-X)) = \sinh(L)\cosh(L-X) - \sinh(L-X)\cosh(L)) \\
 &= \frac{V_L \sinh(X) + V_0 \sinh(L-X)}{\sinh(L)}
 \end{aligned} \tag{4.1.3.2.1}$$

Axial current is,

$$I_i(X) = -\frac{1}{r_a} \frac{\partial V}{\partial X} = -\frac{1}{r_a \lambda} \frac{V_L \cosh(X) - V_0 \cosh(L-X)}{\sinh(L)}$$

Current flowing into the load is, $I_i(X)$ at $X = L$, which is,

$$I_i = -\frac{1}{r_a} \frac{\partial V}{\partial X} = -\frac{1}{r_a \lambda} \frac{V_L \cosh(L) - V_0}{\sinh(L)}$$

Since load current is also equal to, V_L / R_L , we have

$$-\frac{1}{r_a \lambda} \frac{V_L \cosh(L) - V_0}{\sinh(L)} = \frac{V_L}{R_L},$$

or,

$$V_L = \frac{V_0 R_L}{R_\infty \sinh(L) + R_L \cosh(L)}$$

$$= \frac{V_0}{\frac{R_\infty}{R_L} \sinh(L) + \cosh(L)}$$

Substituting the above in eqn. 4.1.3.2.1,

$$\begin{aligned}
 V(X) &= \frac{V_L \sinh(X) + V_0 \sinh(L-X)}{\sinh(L)} \\
 &= \frac{V_0 \sinh(L-X) + \frac{V_0}{\frac{R_\infty}{R_L} \sinh(L) + \cosh(L)} \sinh(X)}{\sinh(L)} \\
 &= V_0 \frac{\left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L) \right) \sinh(L-X) + \sinh(X)}{\sinh(L) \left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L) \right)} \\
 &= V_0 \frac{\frac{R_\infty}{R_L} \sinh(L) \sinh(L-X) + \cosh(L) \sinh(L-X) + \sinh(L-(L-X))}{\sinh(L) \left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L) \right)} \\
 &= V_0 \frac{\frac{R_\infty}{R_L} \sinh(L) \sinh(L-X) + \sinh(L) \cosh(L-X)}{\sinh(L) \left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L) \right)}
 \end{aligned}$$

On further reduction we obtain,

$$V(X) = V_0 \frac{\frac{R_\infty}{R_L} \sinh(L-X) + \cosh(L-X)}{\left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L) \right)}$$

Applying the formula for the input current,

$$I_i = -\frac{1}{r_a} \frac{\partial V}{\partial x} \Big|_{x=0} \quad (4.1.3.2.2)$$

We can obtain the expression for R_{in} as,

$$R_{in} = \frac{V_0}{I_i(X=0)} = R_\infty \frac{R_\infty \tanh(L) + R_L}{R_L \tanh(L) + R_\infty} \quad (4.1.3.2.3)$$

4.1.4 Time-dependent solution:

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} = \tau \frac{\partial V_m}{\partial t} + V_m - I_{inj} r_m$$

Let,

$$T = \frac{t}{\tau_m}; \quad X = \frac{x}{\lambda}$$

$$\frac{\partial^2 V_m}{\partial X^2} = \frac{\partial V_m}{\partial t} + V_m - \frac{I_{inj}(X, T)}{\lambda c_m}$$

$$I_{inj}(X, T) = \lambda \tau_m I_{inj}(x, t) \quad (\text{current density})$$

Infinite cable:

Impulse response

Boundary condition: $V(X) \rightarrow 0$, as $|X| \rightarrow \infty$

$$\text{If, } I_{inj} = I_0 = \frac{Q_0}{\tau_m} \quad (\text{an infinitely brief pulse})$$

$$V(X, T) = \frac{I_0 r_m}{2 \lambda \sqrt{\pi T}} e^{-\frac{X^2}{4T}} e^{-T}$$

For long times, decay pattern is the same throughout the cable.
Since the system is linear, response to arbitrary current injection is given as,

$$\begin{aligned} V(X, T) &= \frac{\tau_m}{Q_0} V_\delta(X, T) * I_{inj}(T) \\ &= \frac{\tau_m}{Q_0} \int V_\delta(X, T') * I_{inj}(T - T') dT' \end{aligned} \quad (4.1.4.1)$$

Voltage response to current step:

$$I_{inj} = I_{step} = I_0 u(t)$$

$$V(X, T) = I_0 \frac{\tau_m}{Q_0} \int_0^T V_\delta(X, T') dT'$$

$$V_{step}(0, T) = \frac{I_0 R_\infty}{2} \operatorname{erf}(\sqrt{T}) \quad (4.1.4.2)$$

Example:

Plot the time dependant solutions for different distances from the point of injection for a semi-infinite cable with respect to time. Current is injected at X=0;

Solution:

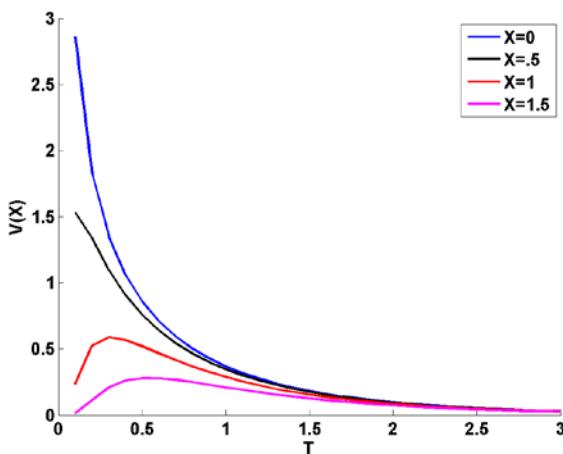
The required plots can be generated using the formula:

$$V(X, T) = \frac{I_0 r_m}{\lambda \sqrt{\pi T}} e^{-\frac{x^2}{4T}} e^{-T}$$

Approximating it as:

$$V(X, T) = B e^{-\frac{x^2}{4T}} e^{-T}$$

We plot the X's(distance's) at intervals of 0.5.



Propagation Delay:

Linear cable equation does not admit any wave equation, due to dissipation of energy through the passive membrane. If voltage-sensitive, nonlinear elements are sensitive, then we can have traveling waves as it happens in the axons.

One way of defining velocity is by tracking motion of the peak.
But a centroid gives uniform propagation velocity.

$$\hat{t}_x^h = \frac{\int_{-\infty}^{\infty} t h(x, t) dt}{\int_{-\infty}^{\infty} h(x, t) dt} \quad (4.1.4.3)$$

Transfer delay:

Transfer delay is the difference in centroid of induced voltage at y, and the centroid of injected current at 'x'.

$$D_{x \rightarrow y} = D_{xy} = \hat{t}_y^V - \hat{t}_x^I \quad (4.1.4.4)$$

Input or local delay:

$$D_{xx} = \hat{t}_x^V - \hat{t}_x^I$$

- D_{xy} is always positive and is independent of the shape of the transient input current.
- It is the property of the cable and not of the input waveform

No matter what the electrical structure of the cable, the transfer delay is symmetric.

$$D_{xy} = D_{yx}$$

It does not depend on the direction of travel.

Isopotential neuron – membrane patch:

$$D_{xx} = \tau_m$$

Infinite or semi-infinite cable:

$$D_{xx} = \frac{\tau_m}{2}$$

If the current is injected at x and the voltage is recorded at y, the two centroids are displaced by,

$$D_{xx} = \left(1 + \frac{|x-y|}{\lambda}\right) \frac{\tau_m}{2}$$

Propagation delay in difference between centroids of voltage at x and y:

$$P_{xy} = \hat{t}_y^V - \hat{t}_x^V = D_{xy} - D_{xx}$$

For an infinite cable,

$$P_{xx} = \left(\frac{|x-y|}{\lambda} \right) \frac{\tau_m}{2}$$

$$v = \frac{\text{distance}}{\text{delay}} = \frac{2\lambda}{\tau_m} = \left(\frac{d}{R_m R_i C_m} \right)^{1/2} \quad \text{"pseudo-velocity"}$$

4.1.5 Branched cables:

Assume that the 2 rightmost branches terminate with sealed end (open) boundary conditions.

Input resistance of the branches,

$$R_{in,1} = R_{\infty,1} \coth(L_1) \tag{4.1.5.1}$$

$$R_{in,2} = R_{\infty,2} \coth(L_2) \tag{4.1.5.2}$$

Terminal resistance of main cable, $R_{L,0}$ is given as,

$$\frac{1}{R_{L,0}} = \frac{1}{R_{in,1}} + \frac{1}{R_{in,2}} \tag{4.1.5.3}$$

Given the terminal resistance, the input resistance is calculated as,

$$R_{in,0} = R_{\infty,0} \frac{R_{L,0} + R_{\infty,0} \tanh(L_0)}{R_{\infty,0} + R_{L,0} \tanh(L_0)} \tag{4.1.5.4}$$

Moving from outermost dendritic tips, one moves towards the soma. Finally at X=0, in the parent cable,

$$V_0 = R_{in,0} I_{inj}$$

Given V_0 , we can use, voltage distribution along the main branch is,

$$V(X) = V_0 \frac{\frac{R_\infty}{R_L} \sinh(L-X) + \cosh(L-X)}{\left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L)\right)} \quad (4.1.5.5)$$

From the above, we can calculate V_{12} , the voltage at the bifurcation point as,

$$V_{12} = V(X)|_{X=L} = \frac{V_0}{\left(\frac{R_\infty}{R_L} \sinh(L) + \cosh(L)\right)}$$

Voltage distribution along the two terminal branches is,

$$V_1(X) = V_{12} \frac{\cosh(L_1 - X)}{\cosh(L_1)}$$

$$V_2(X) = V_{12} \frac{\cosh(L_2 - X)}{\cosh(L_2)}$$

Upward pass (terminals to the root) – compute R_{in} 's::

Compute $R_{in,1}$ and $R_{n,2}$ using eqns. (4.1.5.1, 4.1.5.2). Using $R_{in,1}$ and $R_{n,2}$ compute $R_{L,0}$ using,

Eqn. (4.1.5.3). Using all the 3 results compute, $R_{in,0}$ using eqn. (4.1.5.4).

Downward pass (Root to terminals – compute $V(x)$):

4.1.6 Rall's condition:

When the diameters of the two terminal branches are related to the diameter of the main branch in a special way.

$$d_0^{3/2} = d_1^{3/2} + d_2^{3/2} \quad (4.1.6.1)$$

Lengths of the two terminal branches must be the same i.e., $L_1 = L_2 = L$ (say).

$$R_\infty = (R_m R_i)^{1/2} \frac{2}{\pi d^{3/2}} = \frac{a}{d^{3/2}}$$

Therefore,

$$R_{\infty,0} = \frac{a}{d_0^{3/2}}, R_{\infty,1} = \frac{a}{d_1^{3/2}} \text{ and } R_{\infty,2} = \frac{a}{d_2^{3/2}}$$

Note that, thanks to the way the diameters of the 3 cables are related, we have a special relation among the R_{∞} 's also,

$$\frac{1}{R_{\infty,0}} = \frac{1}{R_{\infty,1}} + \frac{1}{R_{\infty,2}}$$

Input resistance of the terminal branches are,

$$R_{in,1} = R_{\infty,1} \coth(L) \text{ and } R_{in,2} = R_{\infty,2} \coth(L)$$

Thus, terminal resistance of the main cable is given as,

$$\begin{aligned} \frac{1}{R_{L,0}} &= \frac{1}{R_{in,1}} + \frac{1}{R_{in,2}} \\ &= \left(\frac{1}{R_{\infty,1}} + \frac{1}{R_{\infty,2}} \right) \tanh(L) \\ &= \frac{1}{R_{\infty,0}} \tanh(L) \end{aligned}$$

Or,

$$R_{L,0} = R_{\infty,0} \coth(L) \quad (4.1.6.2)$$

From, eqn. (4.1.5.5) we may write down the voltage distribution of the main cable as,

$$V(X) = V_0 \frac{R_{\infty,0} \sinh(L_0 - X) + R_{L,0} \cosh(L_0 - X)}{(R_{\infty,0} \sinh(L_0) + R_{L,0} \cosh(L_0))}$$

Inserting the expression for $R_{L,0}$ from eqn. (4.1.5.3) here,

$$\begin{aligned} V(X) &= V_0 \frac{R_{\infty,0} \sinh(L_0 - X) + R_{\infty,0} \coth(L) \cosh(L_0 - X)}{(R_{\infty,0} \sinh(L_0) + R_{\infty,0} \coth(L) \cosh(L_0))} \\ &= V_0 \frac{\sinh(L) \sinh(L_0 - X) + \cosh(L) \cosh(L_0 - X)}{\sinh(L) \sinh(L_0) + \cosh(L) \cosh(L_0)} \\ &= V_0 \frac{\cosh(L_0 + L - X)}{\cosh(L_0 + L)} \end{aligned}$$

Rall's conditions:

- 1) R_m, R_i are the same in all branches
- 2) All terminals end in the same boundary condition
- 3) All terminal branches end at the same electrotonic distance from the origin of the main branch
- 4) At every branch point, infinite input resistances must be matched, i.e.,

$$d_0^{3/2} = \sum_i d_i^{3/2}$$

Where, d_i are the diameters of the branch cables and d_0 is the diameter of the root cable at the branch cable.

- 5) Identical synaptic inputs must be delivered to all corresponding dendritic locations.

$$X = L_0 + X_1$$

4.2 Axon:

From the Hodgkin-Huxley model (eqns. (3.4.2-3.4.10)), we know the mechanisms underlying membrane excitation. We know how currents that exceed a threshold value can produce action potentials. But the Hodgkin-Huxley model is an isopotential model. It applies to a patch of membrane with homogenous distribution of ion channels and a uniform membrane voltage. Therefore, it does not describe signal propagation. With a slight modification, the Hodgkin-Huxley model can be transformed into a model of voltage wave propagation in an axon.

Eqn. (4.1.1.3) depicts the relationship between the membrane current, I_m , and second spatial derivative of membrane voltage.

$$\frac{1}{r_a} \frac{\partial^2 V_m}{\partial x^2} = I_m(x, t) \quad (4.2.1)$$

This relationship can be used to develop a model of action potential propagation in the axon.

Reproducing eqn. (3.4.2) here with a slight rearrangement of terms,

$$C \frac{dV_m}{dt} + g_{Na}^{\max} m^3 h(V_m - E_{Na}) + g_K^{\max} n^4 (V_m - E_K) + g_l (V_m - E_l) = I_{ext}$$

I_{ext} in the above equation represents I_m in eqn. (4.2.1) expressed per unit length.

Currents in eqn. (3.4.2) are expressed as per unit area, whereas current in eqn. (4.2.1) is expressed in per unit length. To describe propagation along a cable, we need to express all membrane currents as per unit length.

From eqn. (4.1.1.) we know that,

$$r_a = \frac{R_i}{A} = \frac{R_i}{\pi d^2 / 4} = \frac{4R_i}{\pi d^2}$$

For a cable of diameter, d , the membrane current per unit length of the axon is expressed by dividing the currents in eqn. (4.2.1) by $\pi * d * 1$,

$$\frac{1}{\pi dr_a} \frac{\partial^2 V_m}{\partial x^2} = \frac{I_m(x, t)}{\pi d}$$

Or,

$$\frac{d}{4R_i} \frac{\partial^2 V_m}{\partial x^2} = \frac{I_m(x, t)}{\pi d}$$

Substituting the last equation in eqn. (3.4.2),

$$\frac{d}{4R_i} \frac{\partial^2 V_m}{\partial x^2} = C \frac{dV_m}{dt} + g_{Na}^{\max} m^3 h(V_m - E_{Na}) + g_K^{\max} n^4 (V_m - E_K) + g_l (V_m - E_l) \quad (4.2.2)$$

Eqn. (4.2.2) along with eqns. (3.4.3-3.4.10) constitute the complete set of equations that describe signal propagation along an axon.

Eqn. (4.2.2) can be written as,

$$\frac{\partial^2 V_m}{\partial x^2} = \frac{4R_i}{d} \left(C \frac{dV_m}{dt} + g_{Na}^{\max} m^3 h(V_m - E_{Na}) + g_K^{\max} n^4 (V_m - E_K) + g_l (V_m - E_l) \right) \quad (4.2.3.)$$

We need to seek propagating wave solutions for the above equation. Assuming the existence of a voltage wave, $V(x - u_t)$, propagating with a uniform velocity, u , along the axon, the wave must also satisfy the wave equation,

$$\frac{\partial^2 V_m}{\partial x^2} = \frac{1}{u^2} \frac{\partial^2 V_m}{\partial t^2} \quad (4.2.4)$$

Substitution (4.2.4), we may write eqn. (4.2.3) as,

$$\frac{1}{K} \frac{d^2 V_m}{dx^2} = \frac{dV_m}{dt} + \frac{I_{ionic}}{C_m}$$

Where I_{ionic} is the sum of all the ionic currents in eqn. (4.2.3) and,

$$K = \frac{4R_i u^2 C_m}{d}$$

Note that the currents are expressed in terms of per unit area, therefore they are independent of the axon diameter, d . Similarly the voltage V_m must also be independent of d . Therefore K must also be independent of d . Since R_i and C_m are independent of d by definition, it implies that u^2/d is independent of d . In other words,

$$u \propto \sqrt{d} \quad (4.2.5)$$

Thus we have a relationship between cable diameter and conduction velocity. This rule is roughly followed in real, unmyelinated axons.

Note that the result of eqn. (4.2.5) is identical to that of eqn. (4.1.4.5), the derivation is quite different in the two cases.

Just as the Hodgkin-Huxley model, which is an isopotential model, exhibits threshold effect, the model of wave propagation also exhibits a threshold effect. When suprathreshold external current is injected in at a point in an axon, Na^+ ions rush in thereby increasing local membrane voltage rapidly. Voltage increase in neighboring regions activate the local Na^+ channels further amplifying the voltage buildup. The fraction of membrane current that depolarizes neighboring membrane segments is called ‘local circuit current.’ As this process continues, the action potential spreads along the axon. When a subthreshold current is injected, the local Na^+ channels are not activated, and therefore, due to inadequate amplification, there is no local spread of action potential.

When the current is close to the threshold (within 1% of threshold current), there may not be a propagating wave, but a decaying wave that dies down to resting potential with increasing distance from the point of injection.

Another feature that controls action potential propagation consists of the values of channel conductances. Since the action potential generation and propagation is dependent crucially on voltage-sensitive Na⁺ and K⁺ channels, the conductances of these two channels must be sufficiently high for signal propagation. In a simulation study of signal propagation in axon, Cooley and Dodge (1966) reduced g_{Na} and g_K by a scaling factor η. They found that for η <= 0.26, action potential was generated but did no propagate as a stable wave. Only a decaying wave was observed.

4.2.1 Voltage-sensitive ion channels

In the previous chapter, we have seen how we can express channel models in terms of gating variables. Specifically we considered voltage-sensitive Na⁺ and K⁺ channels used in the Hodgkin-Huxley model. Both of these models had a common mathematical structure. Channel current, I_x, is expressed as a function of gating variables, an activation variable, m, and an inactivation variable, h, as follows:

$$I_x = \bar{g}m^p h^q (V_m - E_x) \quad (4.2.1.1)$$

In the HH model, p=3 and q = 1 for the Na⁺ channel, and p=4, q=0 (for the activation variable) for K⁺ channel. Separate α and β functions were associated with each channel. The form of the conductance variation is given in eqn. (4.2.1), and the α and β functions determine the dynamics of the channel.

We now present models of a greater variety of channels. While the above modeling structure applies to most cases, there will be some deviations. Broadly 4 classes of channels will be considered:

1. Sodium channels
2. Potassium Channels
3. Chloride Channels
4. Calcium channels
 - a. Calcium gated potassium channels

4.2.1.1 Sodium channels:

Sodium channels are broadly classified into “transient” and “persistent” type.

The sodium channels of HH model are fast channels in which the conductance rises fast, due to fast activation dynamics, and drops rapidly, due to fast inactivation dynamics. Such currents are usually called “transient” since the channel opens briefly and shuts again, even though the membrane is depolarized. A transient sodium current, like the one in HH model, can be expressed as,

$$I_{Na,t} = \bar{g}m^3 h(V_m - E_{Na}) \quad (4.2.1.1.1)$$

But there are sodium channels which do not have inactivation gates ie., h variable is absent. Such channels open on depolarization and remain open as long as the membrane

remains depolarized. Such sodium channels are said to be “persistent” ($I_{Na,p}$) or “noninactivating” (referring to the absence of inactivation dynamics). $I_{Na,p}$ currents are found for example in the thalamocortical neurons in rats (Parri and Crunelli 1999). A persistent sodium current may be expressed as,

$$I_{Na,p} = \bar{g}m(V_m - E_{Na}) \quad (4.2.1.1.2)$$

4.2.1.2 Potassium channels

Voltage-sensitive potassium channels are broadly classified in terms of 1) the speed of their inactivation dynamics, and 2) the direction of the currents. Potassium currents with slow inactivation dynamics are said to be “delayed” while those with fast inactivation dynamics are said to carry “A type” currents.

Channels that allow currents in only one direction are said to be rectifying. “Outward rectifying” channels are those that allow current in the outward direction, while “inward rectifying” channels are that allow currents in inward direction. Thus there are three important classes of voltage-sensitive potassium currents:

- 1) Delayed rectifier currents: These currents have slow inactivation dynamics and allow currents in the outward direction. (The ‘rectifier’ implicitly means outward rectifying). The potassium currents in HH model are of this type.
- 2) A currents: These currents have fast inactivation dynamics.
- 3) Inward rectifying currents: Those most potassium channels are outward rectifying, there are a special class of potassium channels (sometimes referred to as the “exceptional”) that are inward rectifying. These are denoted by K_{IR} channels. These channels open under conditions of hyperpolarization.

4.2.1.3 Calcium currents

Like the voltage-sensitive sodium and potassium currents visited above, calcium currents are also described in terms of activation, m , and inactivation, h , variables. But the calcium current cannot be calculated using the simple expression of eqn. (4.2.1.1.1) since the calcium concentration inside the neuron is very low ($O(nM)$) and varies rapidly due to calcium flux. Therefore, the Nernst potential formula which depends on assumptions of equilibrium conditions is not valid. The correct formula for calcium current is given as,

$$I_{Ca} = \bar{p}m^p h^q \frac{V_m z^2 F^2}{RT} \frac{[Ca^{2+}]_e e^{-V_m zF/RT} - [Ca^{2+}]_i}{1 - e^{-V_m zF/RT}} \quad (4.2.1.3.1)$$

Where,

$[Ca^{2+}]_e$ - extracellular calcium concentration (usually about 2 mM).

$[Ca^{2+}]_i$ - intracellular calcium concentration (varies but low)

\bar{p} - max. permeability to calcium

Faraday's constant, $F = 9.648 \times 10^4 \text{ C mol}^{-1}$

Ideal gas constant, $R = 8.314 \text{ V C K}^{-1} \text{ mol}^{-1}$

Though the above description is valid for general calcium currents, there are two subcategories of calcium currents which depend on the threshold voltage at which the channels get activated. These are:

1) Low-threshold calcium current (I_T):

The low threshold calcium channels open at a threshold voltage of about -40 mV. That is, these channels open under conditions of hyperpolarization. Hence I_T current is responsible for an interesting phenomenon called *post-inhibitory rebound*. When an inhibitory, hyperpolarizing input is suddenly switched off, a neuron can show a rebound response whereby, the cell might fire a few action potentials, a phenomenon called post-inhibitory rebound. The role of low threshold calcium channels in post-inhibitory rebound may be explained as follows.

For the low threshold calcium channels under conditions of hyperpolarization (created by inhibitory input), the inactivation variable, h , is positive. But the channel is in closed state since the activation variable, m , is 0. But when the hyperpolarizing current is stopped, membrane voltage gradually increases. Therefore, m gradually increases while h decreases. There will be a critical stage at which both m and h are sufficiently positive, when the channels are briefly open, allowing a transient calcium pulse. Such pulses are known as low-threshold calcium spikes. The resulting calcium influx transiently increase membrane voltage which in turn triggers a couple of sodium spikes.

2) High-threshold calcium current (I_L):

The I_L current differs from its low-threshold counterpart only in the threshold voltage at which the channel gets activated. The High-threshold calcium channels, as the name indicates, open only at high levels of membrane depolarization. They are activated even during the action potentials and contribute to changes in membrane voltage. In addition this class of calcium channels have a role in controlling a class of potassium channels.

2a) Calcium-controlled potassium channels:

Intracellular calcium ions are important players in many forms of second-messenger signaling. An instance of such signaling is the role of intracellular calcium in dynamics of a special class of potassium channels, the calcium-controlled potassium channels. Current through these channels is given as,

$$I_C = \bar{g}_C m(V_m - E_K) \quad (4.2.1.3.2)$$

Dynamics of the activation variable, m , is given in the usual form, as:

$$\dot{m} = \alpha m - \beta(1-m)$$

In the channels of HH model, the α and β functions only depend on membrane voltage, V_m . But in the present case, α depends on intracellular calcium, $[Ca^{2+}]_i$, as follows:

$$\alpha = A [Ca^{2+}]_i \exp(V_m / k)$$

$$\beta = B [Ca^{2+}]_i \exp(-V_m / k)$$

where A, B and k are constants. We know that steady-state value of m may be expressed in terms of α and β functions as,

$$m_\infty = \alpha / (\alpha + \beta)$$

Thus increasing $[Ca^{2+}]_i$ can be seen to increase m, and therefore I_C (up to a point of saturation).

4.3 Synapse

We have already seen that synaptic transmission converts a presynaptic electrical event viz., action potential, into a postsynaptic electrical event – the PSP. This change is produced by the action of the neurotransmitter on a postsynaptic receptor which leads to opening of an ion channel on the postsynaptic side. Therefore the simplest form modeling a synapse would be to consider a membrane model of the postsynaptic side and describe the transient change in the conductance of the ion channel involved in the transmission event.

4.3.1 Circuit model of synapse

Accordingly consider a simple circuit model of the postsynaptic membrane in the figure below.

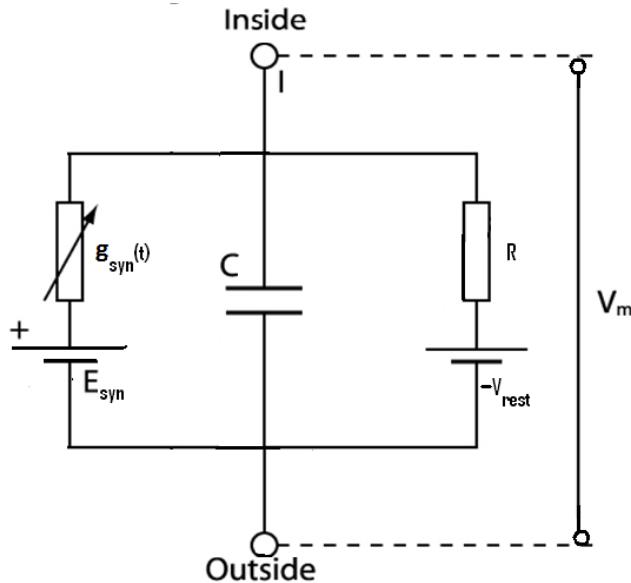


Figure 4.3.1 Simple circuit diagram of Post synaptic membrane

$g_{syn}(t)$ – is the time-varying synaptic conductance (of the ion channel involved in synaptic transmission)

E_{syn} – is the Nernst potential corresponding to the ion channel (involved in synaptic transmission) and the ionic species to which it is permeable

C – membrane capacitance

V_{rest} – resting membrane potential of the postsynaptic membrane

R – membrane resistance

The central and right branches consisting of C, R and V_{rest} together model the passive membrane properties in the absence of synaptic transmission. The left branch becomes active only when there is synaptic transmission, i.e., when $g_{syn}(t) > 0$.

Applying Kirchoff's current law to the above circuit, we have,

$$C \frac{dV_m}{dt} + g_{syn}(t)(V_m - E_{syn}) + \frac{(V_m - V_{rest})}{R} = 0 \quad (4.3.1.1)$$

Note that normal resting conditions of the postsynaptic membrane occur by setting $g_{syn}(t) = 0$ and $V_m = V_{rest}$ in the above equation.

The transient increase and return to zero of synaptic conductance $g_{syn}(t)$ is usually expressed in the following form,

$$g_{syn}(t) = A t e^{-t/t_{peak}} \quad (4.3.1.2)$$

Where A is a constant, and t_{peak} is the time at which $g_{syn}(t)$ peaks. Note that the above expression represents variation of synaptic conductance in response to a single action potential.

Since Post Synaptic Potential (PSP) is defined as the deviation from the resting potential, we define a new voltage variable, v , as,

$$v = V_m - V_{rest}$$

Accordingly eqn. (4.3.1.1) may be rewritten as,

$$C \frac{dv}{dt} + g_{syn}(t)(v + V_{rest} - E_{syn}) + \frac{v}{R} = 0$$

Or,

$$C \frac{dv}{dt} = -(g_{syn}(t) + \frac{1}{R})v + g_{syn}(t)(E_{syn} - V_{rest}) \quad (4.3.1.3)$$

Therefore, during the neurotransmission, when $g_{syn}(t)$ begins to rise, the sign of $(E_{syn} - V_{rest})$ determines the direction of change in v . We have three cases in order.

Case 1: $E_{syn} > V_{rest}$

This amounts to having a positive injected current in eqn. (4.3.1.3). Therefore v transiently increases. This corresponds to an Excitatory Post Synaptic Potential (EPSP).

A common example of synapses that produce such EPSPs is the fast, excitatory synapses with AMPA receptors (or non-NMDA type) and glutamate as neurotransmitter.

Case 2: $E_{syn} < V_{rest}$

Since the current injected is negative, we have a negative deviation in v , which is an Inhibitory Post Synaptic Potential (IPSP).

A common example is the inhibitory synapse with GABA as the neurotransmitter and $GABA_B$ as the receptor. The channel involved is a potassium channel with E_{syn} that is 10-30 mV below the resting potential.

Case 3: $E_{syn} \approx V_{rest}$

There is no current in the synaptic conductance in this case. So it does not seem to have any apparent effect on the postsynaptic potential.

An example of such a synapse is the GABA synapse with $GABA_A$ receptor and chloride as the associated channels. The reversal potential of these channels is close to the resting potential of many cells in which these synapses are found.

Although these synapses do not seem to have any effect in isolation, when used in conjunction with excitatory synapses they show an interesting effect. Let us rewrite eqn (4.3.1.3), so as to include a synapse in which $E_{syn} \approx V_{rest}$, and an excitatory synapse.

$$C \frac{dv}{dt} = -(g_{syn}^0(t) + g_{syn}^e(t) + \frac{1}{R})v + g_{syn}^0(t)(E_{syn}^0 - V_{rest}) + g_{syn}^e(t)(E_{syn}^e - V_{rest}) \quad (4.3.1.4)$$

Where

(g_{syn}^0, E_{syn}^0) - conductance and reversal potential of the synapse where $E_{syn}^0 \approx V_{rest}$

(g_{syn}^e, E_{syn}^e) - conductance and reversal potential of the excitatory synapse

The second term on the RHS is small since $E_{syn}^0 \approx V_{rest}$. But due to the presence of g_{syn}^0 , the conductance term in the first term on RHS is greater (than what it would be when only the excitatory synapse is present). Therefore, under these conditions the EPSP produced is smaller than what it would be when the excitatory synapse alone is present. In that sense, the synapse corresponding to (g_{syn}^0, E_{syn}^0) is inhibiting the excitatory synapse. Therefore it is known as a silent or a shunting inhibition.

4.3.2 Excitatory synapses:

AMPA type synapses: The form of conductance variation of eqn. (4.3.1.2) is applicable for a general synapse. More accurate models have been proposed for specific synapses. A model of synaptic conductance with a double exponential term has been proposed for AMPA synapses (Gabbiani et al 1994):

$$g_{amp} (t) = \bar{g}_{amp} A (e^{-t/\tau_{decay}} - e^{-t/\tau_{rise}}) H(t) \quad (4.3.2.1)$$

\bar{g}_{amp} : maximum value of the synaptic conductance

A : normalizing constant that ensures that the highest value of the bracketed expression is unity

t_{decay}, t_{rise} : decay and rise time constants

H(t): the step function or the Heaviside function

Numerical values of the above parameters used in (Gabbiani 1994) are as follows:

$$\bar{g}_{amp} = 750 \text{ pS}; A = 1.273; t_{decay} = 1.5 \text{ ms}, t_{rise} = 0.09 \text{ ms.}$$

NMDA type synapses: This type of synapses have more complex dynamics than AMPA type synapses, since they are dual-gated: they can be gated by both the neurotransmitter and the membrane voltage. Modeling the conductance change due to the neurotransmitter is similar to the cases seen above. Under conditions of normal membrane polarity, extracellular Mg²⁺ blocks the NMDA associated channels. This block is removed when the membrane potential is depolarized beyond -50 mV. However the time-scales of opening of this channel is longer (10-100 ms) compared to that of AMPA channels (about 1 ms). The form of synaptic conductance for NMDA type synapses is given as follows (Gabbiani et al 1994):

$$g_{nmda}(t) = \bar{g}_{nmda} A(e^{-t/\tau_{decay}} - e^{-t/\tau_{rise}})(1 + e^{\alpha V_m}[Mg^{2+}]_0 / \beta)H(t) \quad (4.3.2.2)$$

Values of various parameters in the last equation are given as (Gabbiani et al 1994),

$$\bar{g}_{nmda} = 1.2 \text{ nS}; A = 1.358; \tau_{decay} = 40 \text{ ms}, \tau_{rise} = 3 \text{ ms}; \alpha = 0.062 \text{ mV}^{-1}; \beta = 3.57 \text{ mM};$$

$$[Mg^{2+}]_0 = 1.2 \text{ mM}.$$

4.3.3 Inhibitory synapses:

Even for inhibitory synapses, there are more complex models of synaptic conductance variation than that given by eqn. (4.3.1.2). For example, in the GABAergic synapses of cerebellar granule cells, postsynaptic current is found to have a fast and a slow component. Synaptic conductance in such a case may be expressed as,

$$g_{gaba}(t) = \bar{g}_{gaba} (\bar{g}_{fast} e^{-t/\tau_{fast}} + \bar{g}_{slow} e^{-t/\tau_{slow}})H(t) \quad (4.3.3.1)$$

References:

- C. Koch, Biophysics of Computation, Oxford University Press, 1999.
- GABBIANI, F., MIDTGAARD, J. & KNOPFEL, T. (1994). Synaptic integration in a model of cerebellar granule cells. Journal of Neurophysiology 72, 999-1009.

5. Simplified Neuron Models

In chapter 3, we have introduced the HH model which is the earliest model of action potential (AP) generation in an axon. It shows how voltage-sensitive dynamics of sodium and potassium channels results in generation of APs. But real neurons have a much large variety of ion channels. There are dozens of voltage- and Ca²⁺-gated channels known today. Combinations of these channels can give rise to an astronomically large number of neuron models. Even the HH model, the simplest model of AP generation that we encountered so far, has 4 differential equations, three of them being nonlinear. More realistic neuron models with larger number of ion channels can easily become mathematically untractable. It would be desirable to construct simplified models, with fewer variables, and milder nonlinearities, in such a way that the reduced model preserves the essential dynamics of their more complex versions. One of the first reduced model of that kind is the Fitzhugh-Nagumo neuron model.

5.1 FitzHugh-Nagumo model

FitzHugh-Nagumo model is a two-variable neuron model, constructed by reducing the 4-variable HH model, by applying suitable assumptions.

Hodgkin- Huxley model:

$$\begin{aligned} C \frac{dv}{dt} + g_{Na} m^3 h (v - E_{Na}) + g_k n^4 (v - E_k) + g_L (v - E_L) &= I_{at} \\ \frac{dm}{dt} &= \alpha_m (V_m) (1 - m) - \beta_m (V_m) m \\ \frac{dh}{dt} &= \alpha_h (V_m) (1 - h) - \beta_h (V_m) h \\ \frac{dn}{dt} &= \alpha_n (V_m) (1 - n) - \beta_n (V_m) n \end{aligned}$$

Assumptions:-

The time scales for m, h and n variables are not all of the same order. These disparities provide a basis for eliminating some of the gating variables.

- 1) Since the time scale for m is much smaller than that of the other two, we assume that m relaxes faster than the other two gating variables. Therefore, we let,

$$\frac{dm}{dt} = 0 \text{ i.e.,}$$

$$\frac{dm}{dt} = \alpha_m(v)(1-m) - \beta_m(v)(m)$$

$$0 = \alpha_m(v) - (\alpha_m(v) + \beta_m(v))(m)$$

$$m = \frac{\alpha_m(v)}{\alpha_m(v) + \beta_m(v)} \quad (5.1.1)$$

2) h varies too slowly. Therefore, we let h to be a constant, $h = h_0$

The resulting system has only two variables (v , w). After transformation to dimensionless variables, and some approximations, the resulting FN model may be defined as,

$$\frac{dv}{dt} = f(v) - w + I_m \quad (5.1.2)$$

$$\text{where } f(v) = v(a - v)(v - 1) \quad (5.1.3)$$

$$\frac{dw}{dt} = b v - r w \quad (5.1.4)$$

In the above system, v is analogous to the membrane voltage, and w represents all the three gating variables.

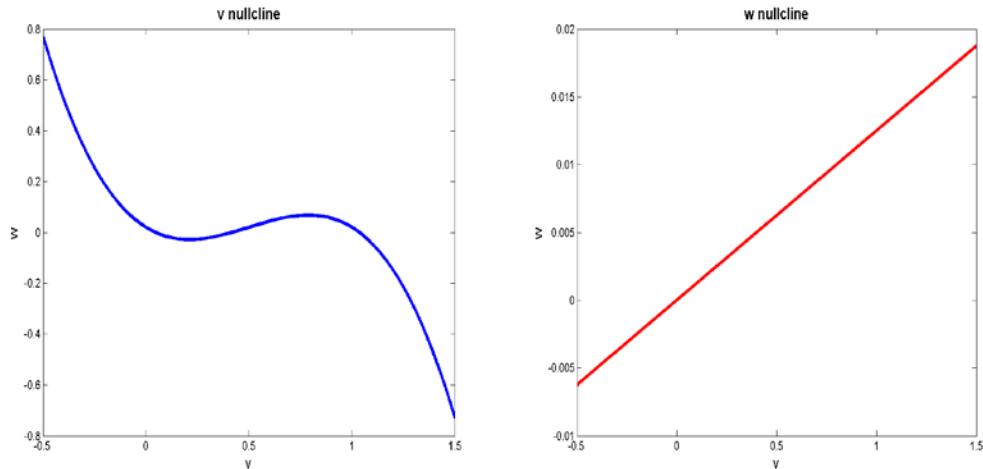


Figure 5.1.1: The nullclines V and w

The two nullclines of the above system are,

$$F(v, w) \equiv f(v) - w + I_a = 0 \quad (\text{F-nullcline})$$

$$g(v, w) \equiv bv - rw = 0 \quad (\text{g-nullcline})$$

To examine the stability of a stationary point, we calculate the Jacobian, A, of the system at that point.

$$A = \begin{bmatrix} \frac{\partial F}{\partial v} & \frac{\partial F}{\partial w} \\ \frac{\partial g}{\partial v} & \frac{\partial g}{\partial w} \end{bmatrix} = \begin{bmatrix} f'(v) & -1 \\ b & -r \end{bmatrix}$$

$$\tau = f'(v) - r \quad (5.1.5)$$

$$\Delta = f'(v)(-r) \quad (5.1.6)$$

The type of the stationary point can be expressed in terms of determinant, Δ , and trace, τ , of the Jacobian, using the following rules:

- if $\Delta < 0$, the stationary point is a saddle irrespective of the value of τ .
- if $\Delta > 0$, $\tau < 0$, stable point
- if $\Delta > 0$, $\tau > 0$, unstable point.

$$\Delta > 0$$

$$f'(v)(-r) > -b$$

$$f'(v)r < b$$

$$f'(v) < \frac{b}{r} \quad (5.1.7)$$

That is, $\Delta > 0$, when the slope of the F-nullcline is lesser than slope of w-nullcline.

$$\tau > 0, \Rightarrow f'(v) - r > 0$$

$$\Rightarrow f'(v) > r \quad (\text{approx}) \quad (5.1.8)$$

Now let us consider the behavior of FN model as external current I_a is gradually increased.

5.1.1) $I_a=0$, Excitability

The phase-plane shown below depicts the situation when $I_a=0$. There is only one stationary point at the origin.

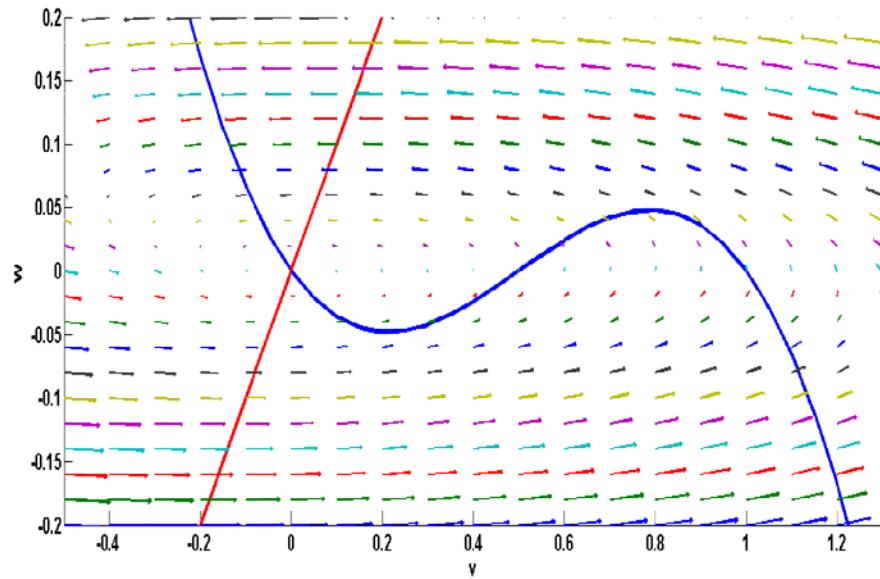


Figure 5.1.1.1: Phase plane analysis: Excitability at $a=0.5$; $b=0.1$; $r=0.1$; $I_a=0$

Stability of origin:

a) Slope of F-nullcline > slope of w-nullcline $\Delta > 0$

b) $f'(v) < 0, \Rightarrow \tau < 0$

\therefore Origin is stable.

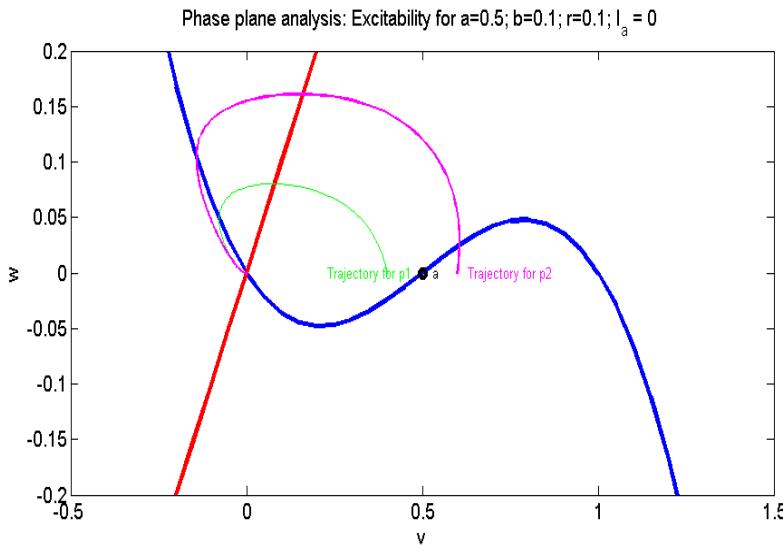


Figure 5.1.1.2: The points a, p_1, p_2 and their trajectories

Consider the evolution of the variable v (membrane voltage) when the initial condition is at points p_1 or p_2 (fig. 5.1.1.2).

This behavior will first be anticipated using loose arguments, and then confirmed using simulations.

The F - and w -nullclines in fig. (5.1.1.1,5.1.1.2) above intersect only at one point (the origin) and therefore divide the plane into four regions, numbered from 1 to 4 (fig. 5.1.1.1). The flow patterns in the four regions can be seen to be as follows:

Region 1: $\dot{v} < 0, \dot{w} > 0$

Region 2: $\dot{v} < 0, \dot{w} < 0$

Region 3: $\dot{v} > 0, \dot{w} < 0$

Region 4: $\dot{v} > 0, \dot{w} > 0$

We have just talked about the signs of \dot{v}, \dot{w} , but it must be noted that, far from the nullclines, the magnitude of \dot{w} is much smaller than that of \dot{v} , since $b, r \ll 1$.

We divide the F -nullcline into three segments: Segment 1 (to the left of point M), Segment 2 (between points M and N) and Segment 3 (to the right of point N). We will refer to these segments in the following discussion.

Let us now consider the two initial conditions:

Initial condition at p_1 : This point is inside region 1. Therefore the flow is leftwards, with a small upward component. The system state approaches the origin and settles there, confirming our earlier result that the origin is the only stable point.

Initial condition at p_2 : This point is inside region 4, where the flow is rightwards with a small upward component. Therefore the system state moves rightwards until it hits the F-nullcline inside Segment 3. Since the flow has no horizontal component on the F-nullcline, the small upward component pushes the state upwards. In Segment 3 of F-nullcline there is a tendency for the state stay on the F-nullcline, since the flow is leftwards on its right, and rightwards on its left. Therefore, the state creeps along the F-nullcline in the upward direction until it reaches the topmost point, N, on the F-nullcline. Beyond this point the flow is still upwards and leftwards while the F-nullcline bends downwards. Therefore, the state leaves the F-nullcline and drifts leftwards until it reaches Segment 1. The situation now similar to what we encountered on Segment 3, but with a downward flow component. Therefore the state creeps downward along Segment 3 until reaches the origin where it finally settles down.

Therefore, when p_2 is the initial condition, the system exhibits this large excursion by which the membrane voltage reaches a maximum before it returns to the origin. Such an excursion of membrane voltage resembles an action potentials.

Therefore the FitzHugh-Nagumo neuron model exhibits excitability. For the initial membrane less than a threshold value ($=a$), the voltage quickly returns to zero (fig. 5.1.1.3). When the initial voltage exceeds the threshold ($=a$), the voltage exhibits an action potential (fig. 5.1.1.3).

$$p_1, p_2 \Rightarrow \text{Initial voltages } v(0)$$

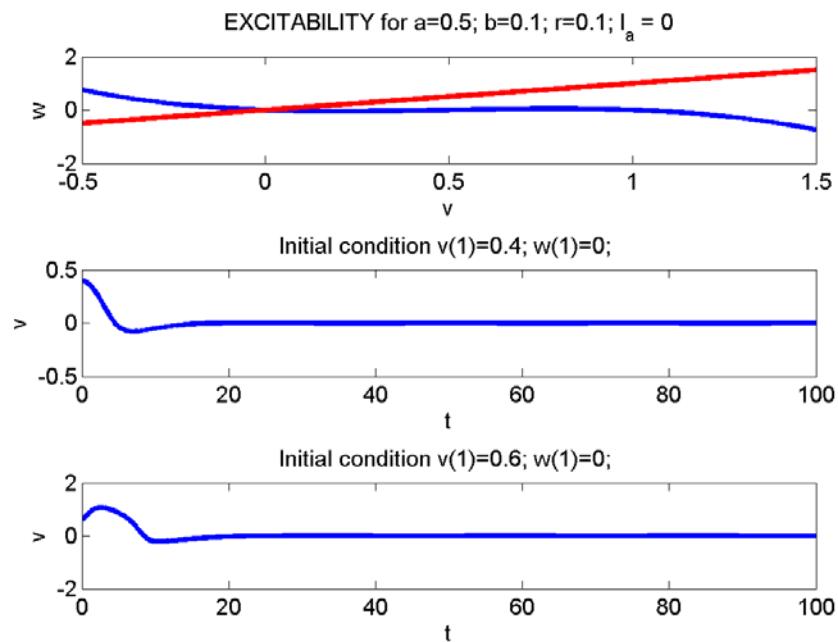


Figure 5.1.1.3: a) The nullclines v and w , and ' v ' simulation from b) p_1 and c) p_2

5.1.2) Limit Cycles ($I_a > 0$):

As I_a increases, for a range of values of I_a , the w-nullcline intersects the F-nullcline in the “middle branch” where the F-nullcline has a positive slope. In this case too there is only a single intersection.

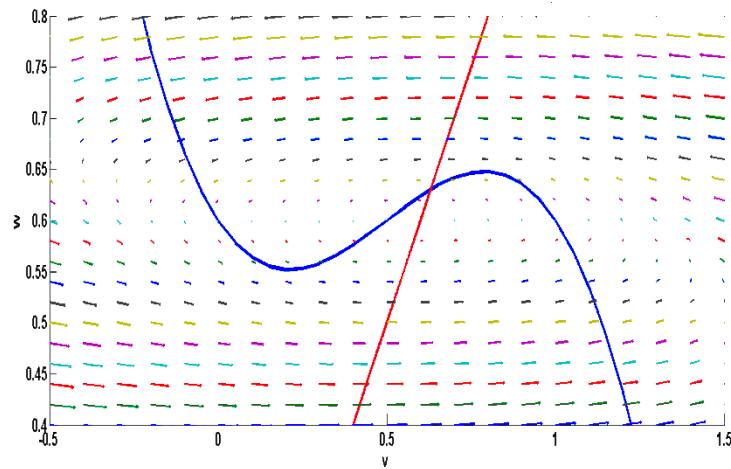


Figure 5.1.2.1: Phase plane analysis: Oscillations at $a=0.5$; $b=0.1$; $r=0.1$; $I_a=0.6$

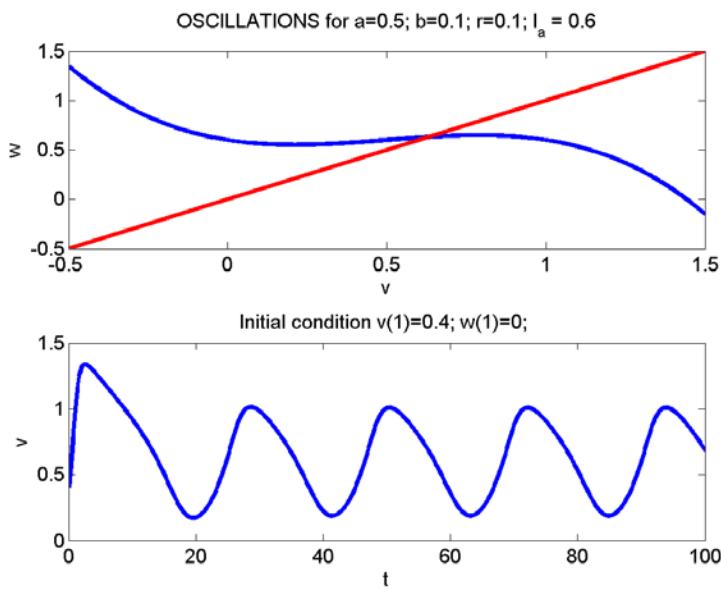


Figure 5.1.2.2: a) The nullclines v and w and 'v' simulation with initial b)
 $v(1)=0.4$

$\tau = f'(v) > 0$ Therefore, the stationary point is unstable.

The rough ‘arrowplot’ in fig.(5.1.2.1) above shows that there is a ‘circulating field’ around the stationary point, which is unstable. Thus it can be expected that there is a limit cycle enclosing the stationary point, which is actually true. Fig. 5.1.2.2 shows the oscillations in membrane voltage (v) produced by a MATLAB program.

5.1.3) Depolarization (higher I_a):

As I_a increases further, the two nullclines intersect in the “right branch” of the F-nullcline where the slope of F-nullcline is negative.

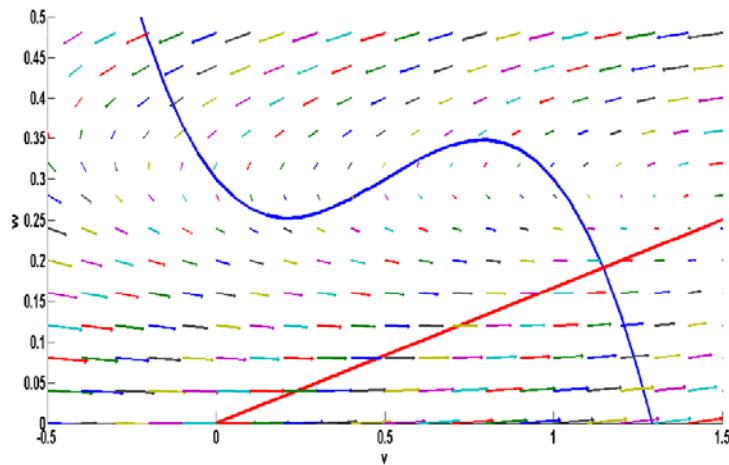


Figure 5.1.3.1: Phase plane analysis: Depolarisation at $a=0.5$; $b=0.1$; $r=0.6$; $I_a=0.3$

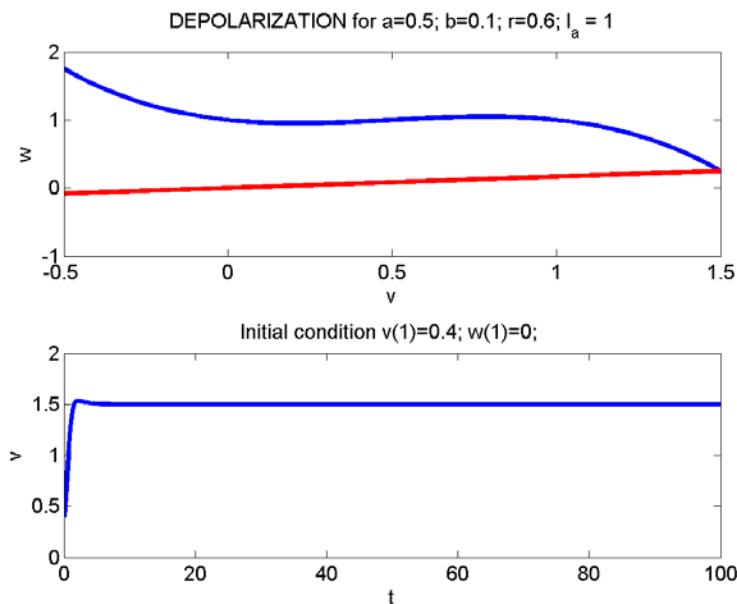


Figure 5.1.3.2: a) The nullclines v and w and 'v' simulation with initial b)
 $v(1)=0.4$

Since,

$$f'(v) < \frac{b}{r}, \Delta > 0$$

$$\tau = f'(v) > 0$$

We know that the stationary point is stable. In this case, the membrane voltage remains stable at a high value. This corresponds to regime 4 (in fig 5.1.3.2) in the HH model where for a sufficiently high current, the neuron does not fire but remains tonically depolarized.

5.1.4) Bistability

Some real neurons exhibit bistable behavior – their membrane voltage can remain at tonically high (“UP” state) or tonically low (“DOWN” state) values. These UP/DOWN neurons are found in for example in medium spiny neurons of Basal ganglia striatum.

The FN model exhibits bistability for a certain range of model parameters.

Fig. 5.1.4.1 below shows a configuration in which the null-clines intersect at three points (p_1 , p_2 and p_3). It can easily be shown that p_1 and p_3 are stable, and p_2 is a saddle.

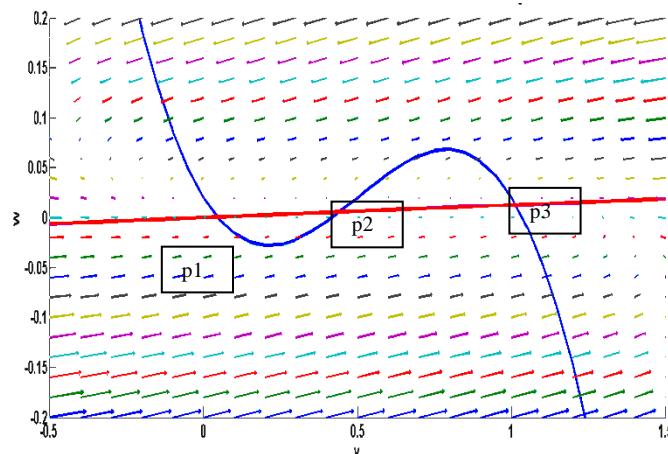


Figure 5.1.4.1: Phase plane analysis: Neuronal on-off bi-stable behavior at $a=0.5$; $b=0.01$; $r=0.8$; $I_a=0.02$

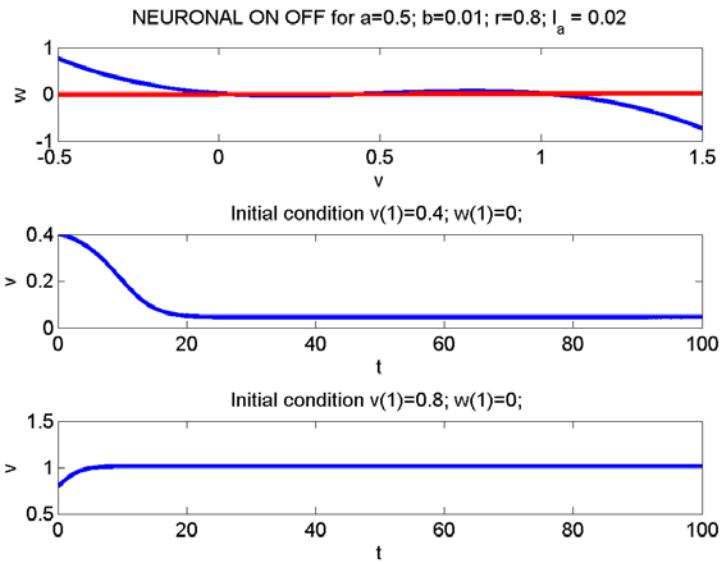


Figure 5.1.4.2: a) The nullclines v and w , and ' v' simulation with initial b)
p1: $v(1)=0.4$ c) p3: $v(1) = 0.8$

$$@ \text{p1: } (V(1) = 0.4) \quad f'(v) < 0 < \frac{b}{r}, \Delta > 0 \\ \tau = f'(v) < 0, \text{ stable}$$

$$@ \text{p2: } (V(1) = 0.5) \quad f'(v) > \frac{b}{r}, \Delta < 0 \\ \tau = f'(v) > 0, \text{ saddle node}$$

$$@ \text{p3: } (V(1) = 0.8) \quad f'(v) < 0 < \frac{b}{r}, \Delta > 0 \\ \tau = f'(v) < 0, \text{ stable}$$

FN model in this case, can remain stable at either p1(low v value, DOWN state) or at p3 (high v value, UP state).

5.2 Morris-Lecar Model

The Morris-Lecar (ML) model (Morris and Lecar 1981) describes the membrane voltage dynamics of the barnacle muscle fiber. It consists of three channel currents:

- a) A fast activating Ca^{2+} current (activating variable – m)
- b) A delayed rectifying K^+ current (activating variable – w)

c) A leakage current

Simplifying Assumption:

As in the case of FN model, we assume that the dynamics of m-variable is fast.
Therefore, m may be substituted by m_∞ .

Equations of the ML model are given as,

$$C \frac{dv}{dt} + g_{Ca} m_\infty (v - E_{Ca}) + g_k w (v - E_k) + g_L (v - E_L) = I_a \quad (5.2.1)$$

$$\tau \frac{dw}{dt} = \phi(-w + w_\infty) \quad (5.2.2)$$

$$m_\infty = \frac{1}{2} [1 + \tanh((v - v_1) / v_2)] \quad (5.2.3)$$

$$w_\infty = \frac{1}{2} [1 + \tanh((v - v_3) / v_4)] \quad (5.2.4)$$

$$\tau = 1 / \cosh((v - v_3) / (2v_4)) \quad (5.2.5)$$

v-nullcline:

$$w = \frac{I_a - g_{Ca} m_\infty (v - E_{Ca}) - g_L (v - E_L)}{g_k w (v - E_k)} \quad (5.2.6)$$

w-nullcline:

$$w = w_\infty = \frac{1}{2} [1 + \tanh((v - v_3) / v_4)] \quad (5.2.7)$$

Note that the v-nullcline is ‘inverted N’-shaped, similar to the F-nullcline of FN model. The w-nullcline has a sigmoidal shape. The v-nullcline can be divided into ‘left’, ‘middle’ and ‘right’ branches. Depending on the branch in which the intersection of V- and w-nullclines, we have different dynamics. The v-nullcline rises with increasing I_a (fig. 5.2.2).

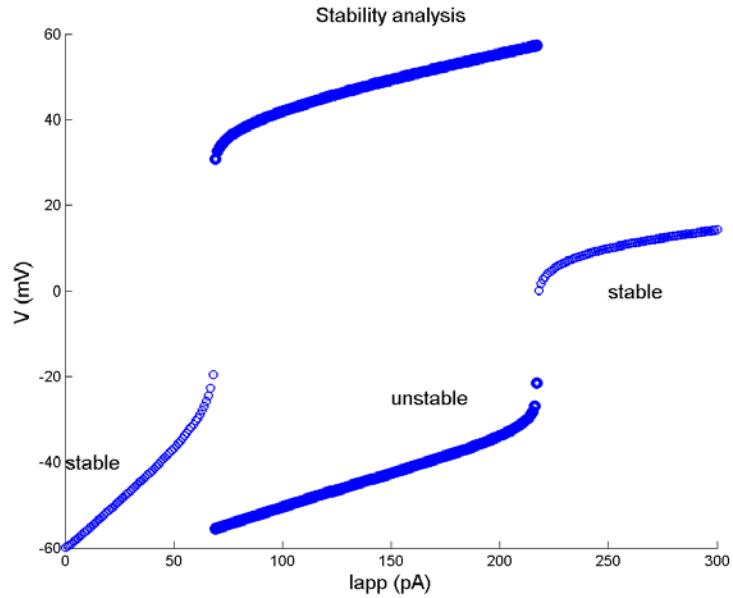


Figure 5.2.1: Bifurcation plot for the ML model

For I_a in the range (0-68 pA): the intersection point is in the left branch. The point can be shown to be stable.

Therefore, the ML model does not produce spikes (fig. 5.2.2).

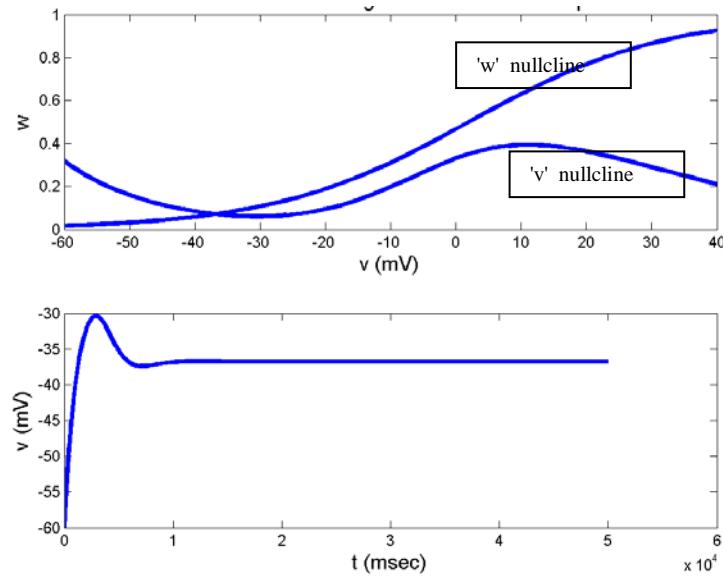


Figure 5.2.2: Nullclines and Voltage simulation for $I = 60$ pA

For I_a in the range (69-217 pA): the intersection point is in the middle branch. Therefore, the ML model produces limit cycle oscillations (fig. 5.2.3).

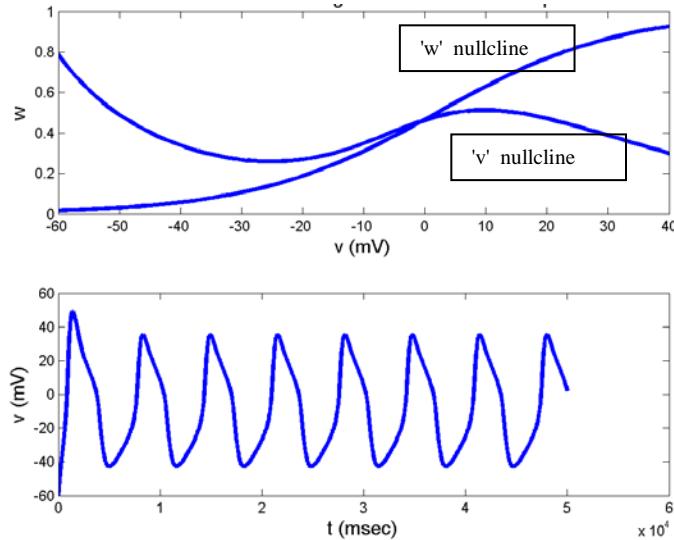


Figure 5.2.3: Nullclines and Voltage simulation for $I = 150$ pA

For I_a in the range (218 pA): the intersection point is in the right branch. Therefore, the membrane voltage exhibits depolarization (fig. 5.2.4).

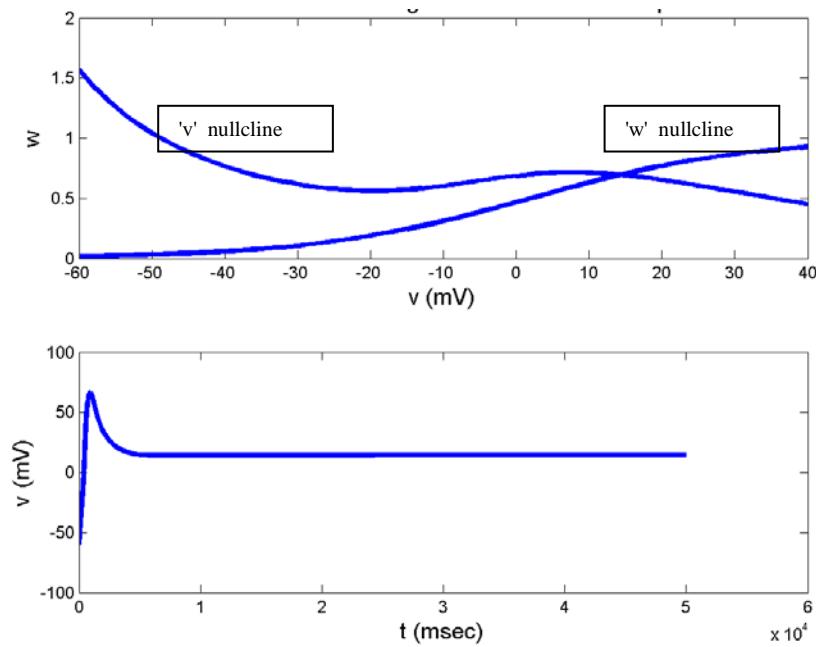


Figure 5.2.4: Nullclines and Voltage simulation for $I = 300 \text{ pA}$

5.3 I_{NaPK} Model

So far we have visited two simplified neuron models that are capable of exhibiting some of the basic dynamic properties of a neuron: resting state, oscillations and depolarization. Both the models are 2-variable systems which are preferred since they can be studied conveniently using phase-plane techniques. The FitzHugh-Nagumo model was reduced systematically and derived from the HH model, whereas the Morris-Lecas model is slightly different from the HH model.

Let us now consider another model that is constructed by systematic reduction from the HH model. This reduced model, known as the $I_{\text{Na},p} + \text{IK}$ model has a persistent sodium channel and a potassium channel. The “persistent” sodium channel (if you recall sodium channel model from chapter 3), has only an activation variable; there is no inactivation variable. The potassium channel also has a single activation variable.

Membrane voltage dynamics of such a model can be written as usual in the following form:

$$C \frac{dv}{dt} = I_a - g_{\text{Na}} m(v - E_{\text{Na}}) - g_k n(v - E_k) - g_L(v - E_L) \quad (5.3.1)$$

$$\tau_m \frac{dm}{dt} = (-m + m_\infty) \quad (5.3.2)$$

$$\tau_n \frac{dn}{dt} = (-n + n_\infty) \quad (5.3.3)$$

If we assume that ‘m’ dynamics is fast, we replace m in eqn. (5.3.1) above with m_∞ and eliminate the ‘m’ dynamics (eqn. (5.3.2)). We are now left with the following two equations:

$$C \frac{dv}{dt} = I_a - g_{\text{Na}} m_\infty(v - E_{\text{Na}}) - g_k n(v - E_k) - g_L(v - E_L) \quad (5.3.4)$$

$$\tau_n \frac{dn}{dt} = (-n + n_\infty) \quad (5.3.5)$$

$$C = 1; g_{\text{Na}} = 20; E_{\text{Na}} = 60 \text{ mV}; g_k = 10; E_k = -90 \text{ mV}; g_L = 8; E_L = -80 \text{ mV}$$

m_∞ and τ_m are modeled on the lines of eqns. (5.3.6,5.3.7).

$$m_\infty = \frac{1}{1 + \exp[(V_{1/2} - V) / \lambda]} \quad (5.3.6)$$

$$\tau_m(V) = C_{base} + C_{amp} \exp[-(V_{max} - V)^2 / \sigma^2] \quad (5.3.7)$$

Thus, the parameters for m_∞ are:

$$V_{1/2} = -20 \text{ mV} \text{ and } \lambda = 15,$$

And the parameters for n_∞ are,

$$V_{1/2} = -25 \text{ mV} \text{ and } \lambda = 5. \tau_n(V) = 1 \text{ for all } V.$$

Rewriting eqns. (5.3.4, 5.3.5) above, we may express the V- and n- null-clines as follows:

V-nullcline:

$$n = \frac{I_a - g_{Na} m_\infty (v - E_{Na}) - g_L (v - E_L)}{g_k (v - E_k)} \quad (5.3.8)$$

and

n-nullcline:

$$n = n_\infty(V) \quad (5.3.9)$$

The V- and n-nullclines obtained from the above equations are shown in fig. 5.3.1, 5.3.2 for $I_a = 60 \text{ pA}$ and 500 pA .

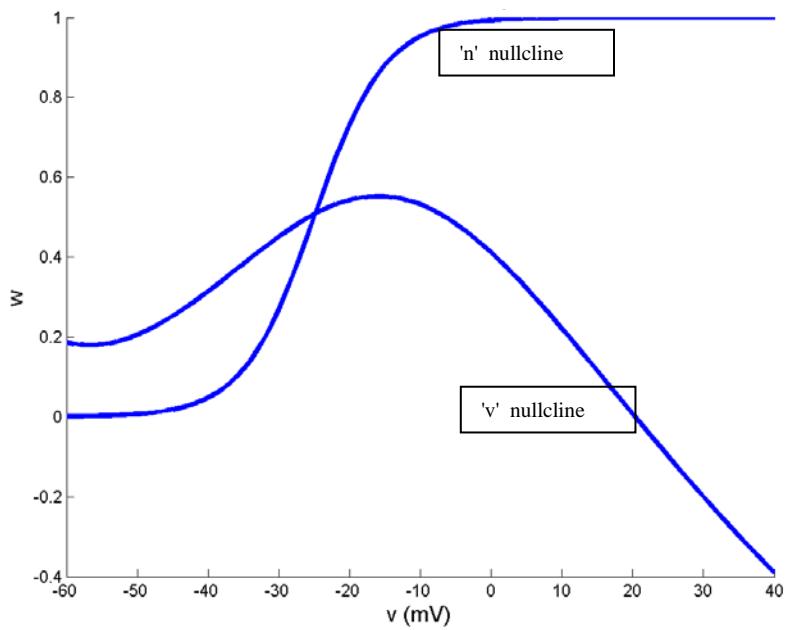


Figure 5.3.1: Nullclines $I_a = 60$ pA

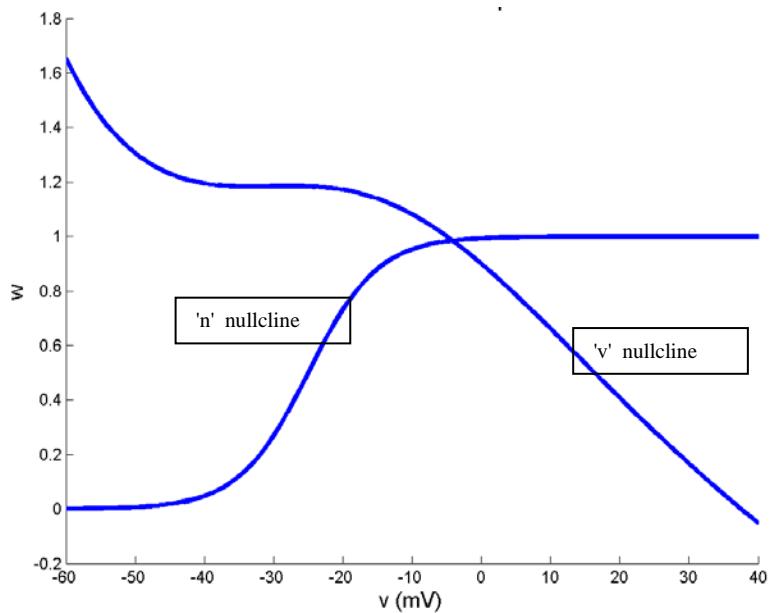


Figure 5.3.2: Nullclines $I_a = 500$ pA

5.4 A Simplified two-dimensional model:

The three models visited above in this chapter have certain common features. They differ in the ion channel composition; they are derived from different cell models by different steps of simplification. But all three share certain features of dynamics. All the three models have:

- Resting state for $I = 0$
- Spikes/oscillations/limit cycles for a sufficiently large I .
- An N-shaped V-nullcline. The shape of the second nullcline is linear in the FN model, and sigmoidal in ML and INAPK model.
- The intersection between the V-nullcline and the second nullcline near the U-shaped, lower arch of the V-nullcline plays a crucial role in neuron dynamics.

When the second nullcline intersects the V-nullcline in the 1st branch there is a resting state; when the intersection occurs in the middle branch there are oscillations.

Consider the role of the remaining part of the state space. In the FN model for example, when the external current $I = 0$. When the initial voltage ($V(0)$) is less than ‘a’, the system returns to the resting potential, 0. But if $V(0) > a$, the neuron state (V, w) increases until it touches the 3rd branch of the V-nullcline, climbs up towards the maximum of the V-nullcline, turns leftwards at the maximum, proceeds up to the 1st branch of the V-nullcline, before it slides down the 3rd branch to the resting state. Thus the remaining part of the phase-plane determines the downstroke and the peak of the action potential.

When the intersection point shifts slightly from the left of the minimum of V-nullcline (where the system exhibits excitability) to the right of the minimum (where the system oscillates), the shape and size of the action potential are about the same. Thus in the oscillatory regime too, the part of the phase-space other than the shaded portion shown in fig. 5.4.1 merely contributes to downstroke and the peak value of the action potential.

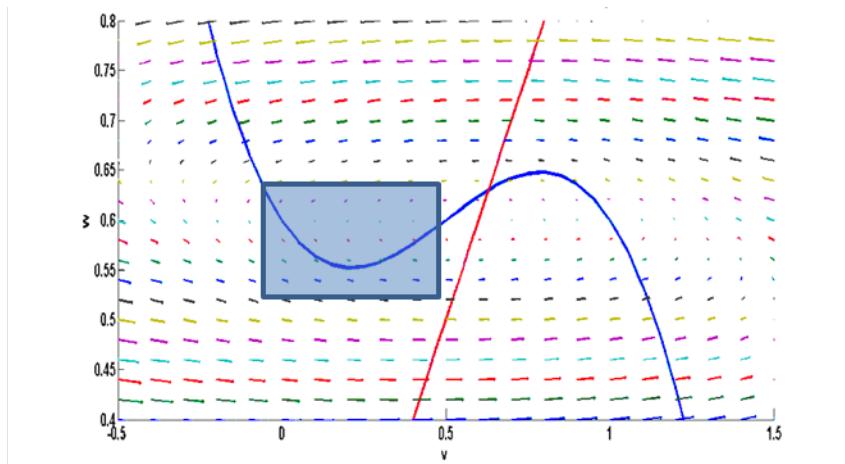


Figure 5.4.1 : Area of interest for the upstroke: Phase plane analysis- Oscillations (FN model) at $a=0.5$; $b = 0.1$; $r=0.1$; $I_a=0.6$

Similar comments can be said about the dynamics in the other two models also.

Thus the above three systems can be reduced to a more general, simpler system as follows. The U-shaped, lower arch of the V-nullcline is approximated by a quadratic function. Since only the shape of the second nullcline at its intersection with V-nullcline matters, and not its shape elsewhere, the second nullcline is approximated by a straightline. Dynamics far from the intersection point are implemented by a simple resetting once the membrane voltage hits a peak value or a minimum.

The quadratic approximation of the V-nullcline at its minimum is given as,

$$u = u_{\min} + p(V - V_{\min})^2 \quad (5.4.1)$$

The linear approximation of the second nullcline may be expressed as,

$$u = s(V - V_0) \quad (5.4.2)$$

The dynamics of the reduced system then becomes,

$$\tau_V \frac{dV}{dt} = -u + u_{\min} + p(V - V_{\min})^2 \quad (5.4.3)$$

$$\tau_u \frac{du}{dt} = -u + s(V - V_0) \quad (5.4.4)$$

Where τ_V and τ_u denote the time-scales of V- and u-dynamics. Since V and u denote the fast and slow variables respectively, $\tau_V \ll \tau_u$.

If the V-dynamics of Eqn. (5.4.3) above is implemented without any auxiliary conditions, V blows up to infinity in a finite time 't'. This can be shown easily on integrating Eqn. (5.4.3) as:

$$V = c_1 * \tan(c_2 t)) \quad (5.4.5)$$

Therefore, the downstroke of the action potential is modeled by resetting $V(t)$ when it reaches the peak value V_{\max} as follows.

$$(V, u) \longleftarrow (V_{reset}, u + u_{reset}), \text{ when } V = V_{\max}. \quad (5.4.6)$$

When the membrane voltage exceeds V_{\max} , both V and u are reset instantaneously as specified by the above equation.

Eqns. (5.4.3, 5.4.4) may be rewritten in a simpler form as,

$$\frac{dv}{dt} = I + v^2 - u \quad (5.4.7)$$

$$\frac{du}{dt} = a(bv - u) \quad (5.4.8)$$

$$\text{If } v \geq 1, v \leftarrow c, u \leftarrow u + d \quad (5.4.9)$$

$a, b, c, d = \text{constants};$

The last set of eqns. (5.4.7- 5.4.9) is generally referred to as the Izhikevich neuron model in current literature. Its merit lies in the low computational cost, and the ability to reproduce firing patterns of a large variety of neurons (E.M. Izhikevich et al, 2004, 2005).

5.4.1 Quadratic integrate and fire neuron:

There is a simpler, one-dimensional version of the Izhikevich model of eqns. (5.4.6- 5.4.8). This model, known as the quadratic integrate and fire neuron model (Latham et al., 2000) consists of only the membrane voltage dynamics with quadratic nonlinearity:

$$\frac{dV}{dt} = I + V^2 \quad (5.4.1.1)$$

$$\text{If } V \geq V_{peak}, V \leftarrow V_{reset} \quad (5.4.1.2)$$

Note that for any non-zero value of I or $V(0)$, eqns. (5.4.1.1) above blows up in a finite time, as already shown above on Eqn. 5.4.5. The resetting condition of eqn. (5.4.6) prevents the blowing up.

By appropriate choice of the model parameters - c - the above model can be made to express a variety of neurodynamic behaviors.

If $I < 0$, $\dot{V} = 0$ at two values, $V = \pm\sqrt{-I}$. Let us call these roots, $V_{threshold} = \sqrt{-I}$ and $V_{rest} = -\sqrt{-I}$. These names can be easily justified.

For $V > V_{threshold}$, $\dot{V} > 0$. Therefore, V increases indefinitely.

Similarly, $V_{rest} < V < V_{threshold}$, $\dot{V} < 0$ and therefore V decreases towards V_{rest} . Therefore $V = V_{threshold}$ is an unstable point.

For $V_{rest} > V$, $\dot{V} > 0$. Therefore, V increases towards V_{rest} . Hence $V = V_{rest}$ is a stable state.

Now consider the dynamics of the neurons in the following 3 cases:

Case i:

For $I < 0$, $V_{threshold} = \sqrt{-I}$ and $V_{rest} = \sqrt{-I}$ are real numbers.

If $V(0) < V_{threshold}$, V approaches V_{rest} and remain there forever.

If $V(0) > V_{threshold}$, V grows indefinitely and, unless clamped, reaches infinity in a finite time.

The reset condition of eqn. (5.4.1.2) prevents the runaway of V , and resets it to V_{reset} as soon as V reaches V_{peak} .

Since $V_{reset} < V_{rest}$, V now tends to V_{rest} and settles there. This latter behavior is described as ‘excitability.’ It is analogous to the case of FN model when $I = 0$.

Excitable behavior may be produced not just by the initial condition, but also by giving a series of pulses which progressively push the membrane voltage towards $V_{threshold}$ and beyond, causing excitation.

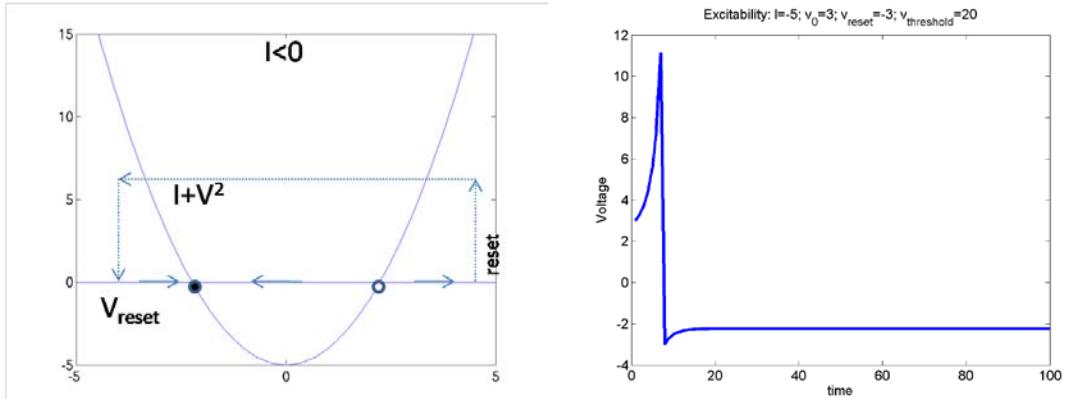


Figure 5.4.1.1: a) Reset condition b) Voltage simulation at $I = 5$; $v_0 = 3$; $v_{reset} = -3$; $v_{threshold} = 20$;

Case (ii):

In this case, the model behavior for $V(0) < V_{threshold}$ is similar to that of the previous case. But when $V(0) > V_{threshold}$, V quickly reaches V_{peak} and gets reset to V_{reset} . But instead of a slow return to V_{rest} , it rises again and again to V_{peak} , exhibiting continuous, periodic firing. Thus the model in case (ii) exhibits two stable states: one corresponding to the resting state, and the other the state of continuous firing. The model is therefore said to have bistability.

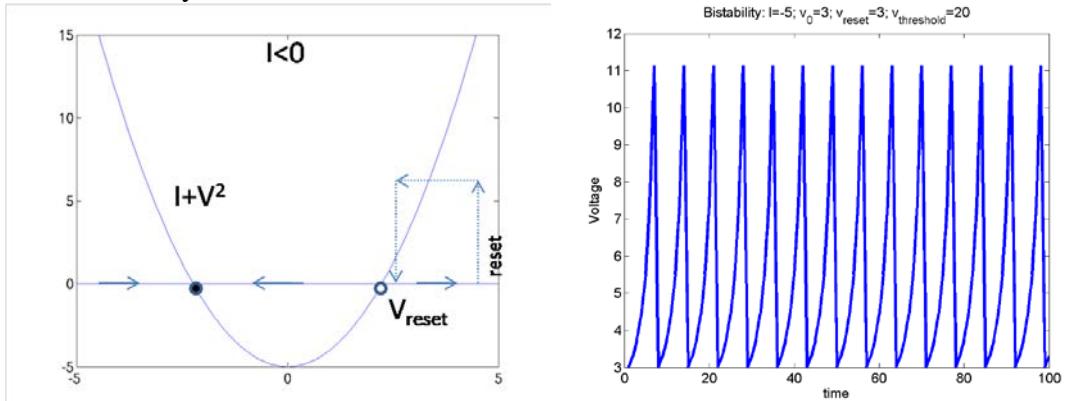


Figure 5.4.1.2: a) Reset condition b) Voltage simulation at $I = -5$; $v_0 = 3$; $v_{reset} = 3$; $v_{threshold} = 20$;

Case (iii): $I > 0$

When $I > 0$ and $V_{threshold} = \sqrt{-I}$ and $V_{rest} = \sqrt{-I}$ are unreal (Fig. 5.4.1.3a). Therefore V rapidly approaches V_{peak} and gets reset to V_{rest} , again and again.

This neuron shows tonic firing, even without excitation.

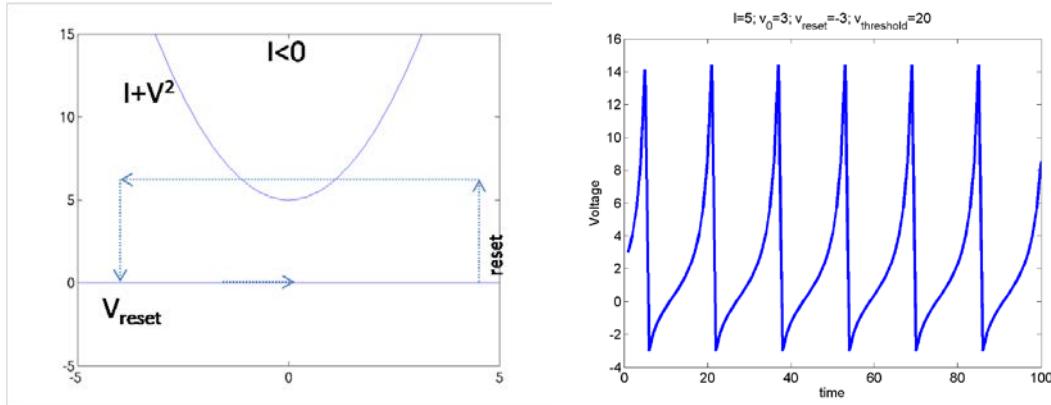


Figure 5.4.1.3: a) Reset condition b) Voltage simulation at $I = 5$; $v_0 = 3$;

5.4.2 Leaky Integrate and fire neuron model:

This is the simplest model of spike generation. An electric circuit implementation of it consists of a capacitance charged by a current, and discharged whenever the voltage across the capacitance exceeds a limit.

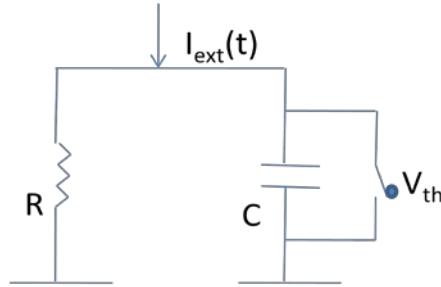


Figure 5.4.2.1: Leaky Integrate and Fire circuit diagram

Applying the Kirchoff's current law of electric circuits, the external current, I_{ext} , going into the circuit may be expressed as,

$$C \frac{dV}{dt} + \frac{V}{R} = I_{ext}(t) \quad (5.4.2.1)$$

Capacitance is discharged whenever, $V > \theta$, a threshold value. It is assumed that whenever the capacitor is charged, the “neuron” emits a spike. Note that in this model, there is no nonlinear, explosive build of excitation reaching a peak producing an action potential. The spike generated in this neuron model is more notional. This has always been one of the points of criticism about the leaky integrate and fire model.

If the capacitance starts off at 0 voltage, let us consider the time taken by the capacitance to reach the threshold, θ .

If I_{ext} is a constant current, I_0 , voltage variation while the capacitance is charging may be expressed as,

$$V(t) = RI_0(1 - \exp(-t / \tau))$$

where τ the time-constant of the circuit equals RC . Since the charging stops at $V = \theta$, the time taken, T , to reach this threshold is given by setting $V(t) = \theta$, or,

$$\theta = RI_0(1 - \exp(-T / \tau))$$

Or

$$T = \tau \ln(RI_0 / (RI_0 - \theta))$$

Since there is a spike every time the capacitance discharges, spike frequency, f , is the reciprocal of T .

$$f = 1 / \tau \ln(RI_0 / (RI_0 - \theta)) \quad (5.4.2.2)$$

A key property of a real neuron reproduced by the above model is thresholding effect. Note that the model exhibits firing only when $R I_0 > \theta$. But as I_0 increases beyond R/θ , f increases indefinitely (Fig. 5.4.2.2), instead of saturating as it happens in a real neuron.

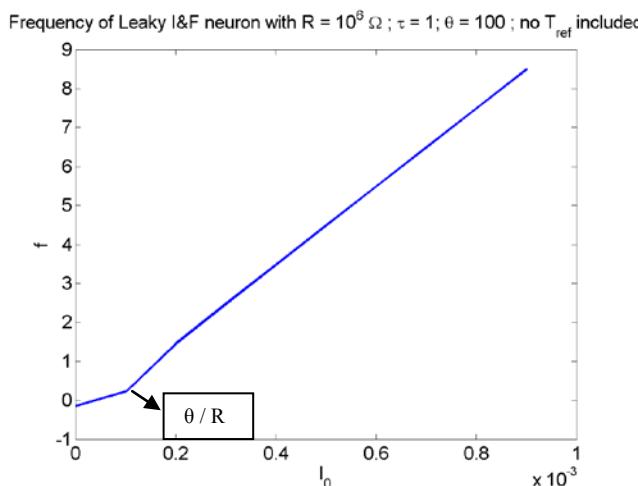


Figure 5.4.2.2: frequency vs I_0 without inclusion of absolute refractory period

In order to restore the saturation property, the condition of absolute refractory period is introduced into the above model. Accordingly the capacitance can begin to be charged

only a little while, T_{ref} , after the external current is applied. Therefore, the time taken to charge is incremented as,

$$T = \tau \ln(RI_0 / (RI_0 - \theta)) + T_{ref}$$

And the new firing rate is,

$$f = \frac{1}{\tau \ln(RI_0 / (RI_0 - \theta)) + T_{ref}} \quad (5.4.2.3)$$

With the inclusion of absolute refractory period, the plot of I_0 vs f , shows saturating behavior (Fig. 5.4.2.3).

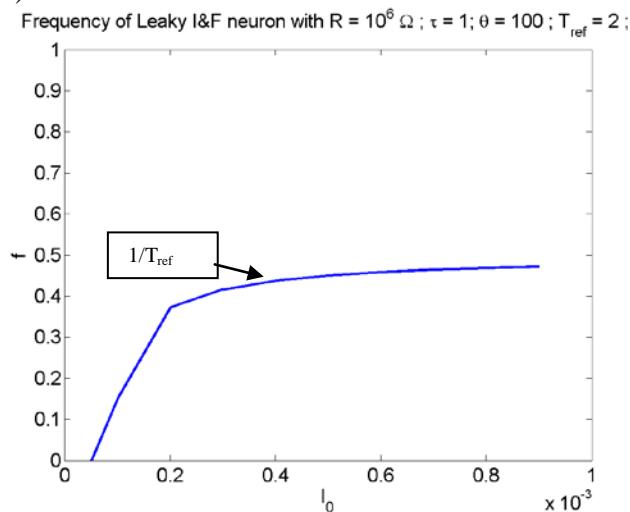


Figure 5.4.2.3: frequency vs I_0 with inclusion of absolute refractory period

5.4.3 Binary neuron models:

The simplified neuron models visited so far in this chapter are models of action potential or spike generation. If we wish to simplify the single neuron model further, we may consider rate-coded models which represent the neuron state in terms of the spike rate. These models typically describe neurons as binary elements, with a high (excited state) and a low (resting state).

5.4.3.1 Dynamic Binary neuron model:

A dynamic version of a binary neuron is a bistable neuron whose dynamics is described as,

$$\tau \dot{u} = -u + V + I \quad (5.4.3.1.1)$$

$$V = \tanh(\lambda u) \quad (5.4.3.1.2)$$

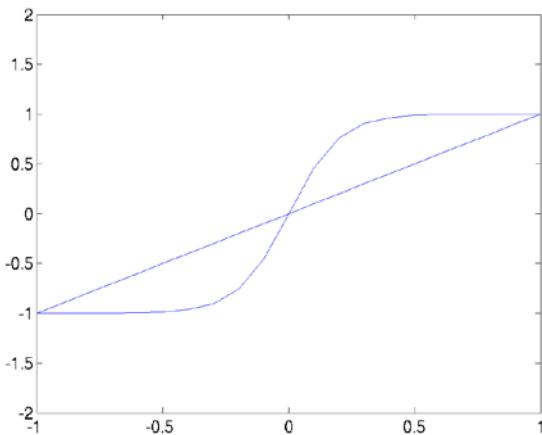


Figure 5.4.3.1.1: u,v nullclines

V is an abstract quantity that denotes if the neuron is excited (V is close to 1) or in the resting state (V is close to -1).

For $I = 0$ and $\lambda > 1$, the above system has three fixed points (Fig. 5.4.3.1.1). The fixed point at the origin can be shown to be unstable, while the ones on the flanks can be shown to be stable. When released from a random initial condition, the neuron state, V, settles near +1 or -1.

For $I > b$, where $b = 2.63$, and $\lambda > 1$, Eqn (5.4.3.1.1) above has only a single, stable fixed point at V close to +1.

For $I < -b$, and $\lambda > 1$, eqn (5.4.3.1.1) above has only a single, stable fixed point at V close to -1.

For $-b < I < b$, there are again three fixed points, two close to +1 and -1 respectively, and the third somewhere in the middle, not necessarily at the origin.

The above neuron model also shows hysteresis effect. If I is reduced from a large positive value to a large negative value and back, the points at which V makes transitions from (+1 to -1 and vice versa) are different in the forward pass and the reverse pass (Fig. 5.4.3.1.2).

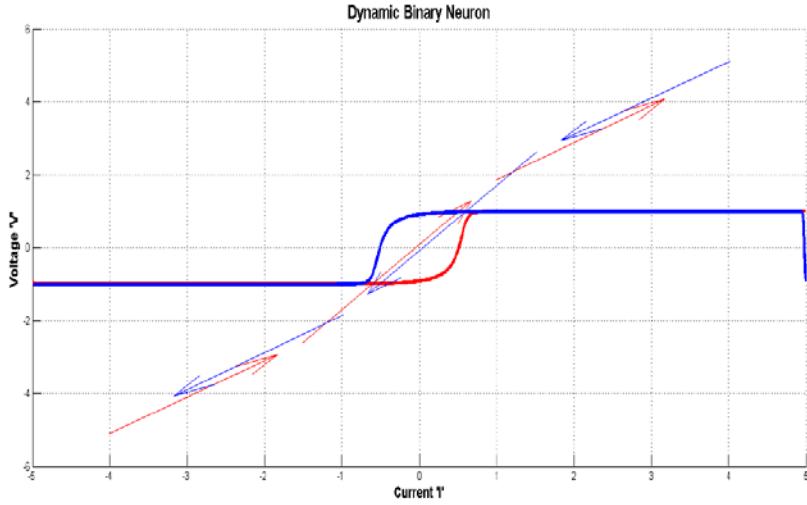


Figure 5.4.3.1.2 : Hysteresis curve for a dynamic binary neuron

5.4.3.2 Static Binary Neuron Model:

The McCulloch and Pitts neuron model is a good example of a static binary neuron. It combines the inputs, x_i , that it receives from other neurons and computes its output, y . The relationship between the inputs x_i and the output y is given as,

$$y = g\left(\sum_{i=1}^n w_i x_i - \theta\right)$$

Where w_i s are the synaptic strengths, or the “weights” of the connections from the neurons that send inputs to the neuron of interest; θ is the threshold for excitation; $g(.)$ is known as a transfer function, which typically has a sigmoidal shape.

Four types of transfer functions are usually considered depending on the range of values that y is permitted.

Hardlimiting nonlinearity:

$$y = 1/0:$$

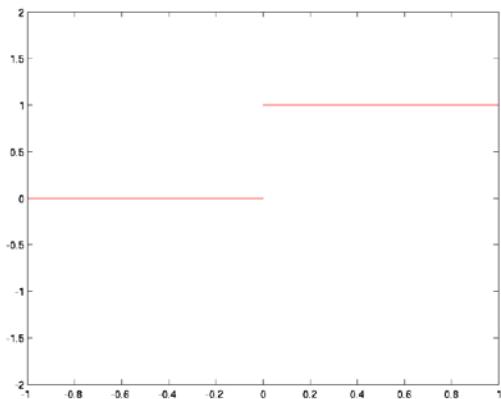


Figure 5.4.3.2.1: Hardlimiting nonlinearity- $y = 1/0$:

$y = +1$.

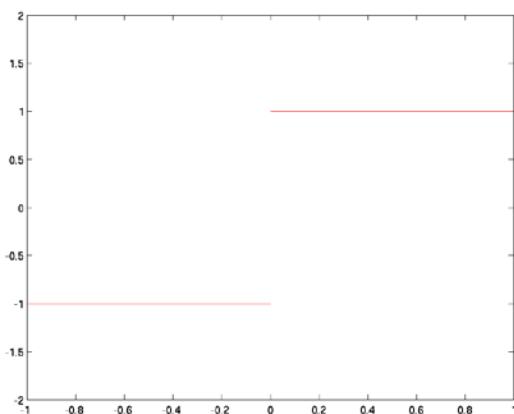


Figure 5.4.3.2.2: Hardlimiting nonlinearity- $y = +1$

Smooth sigmoid nonlinearity:

Logistic function:

$$y \in (0,1)$$

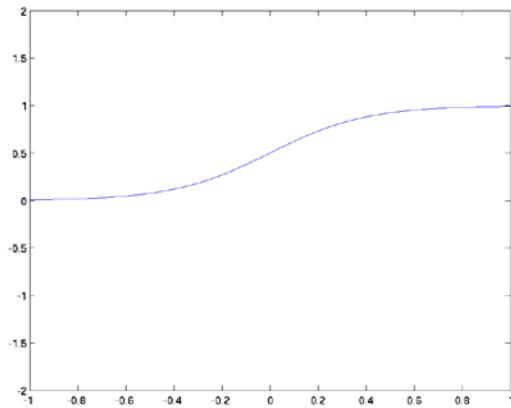


Figure 5.4.3.2.3: Logistic function- $y \in (0,1)$

Tanh(.) function:

$$y \in (-1,1)$$

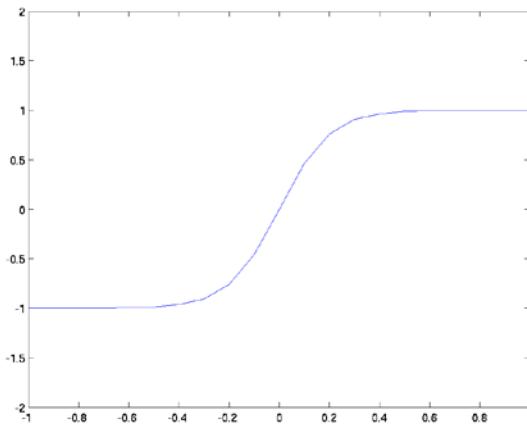


Figure 5.4.3.2.4: Tanh(.) function - $y \in (-1,1)$

References:

E. M. Izhikevich, J. A. Gally, and G. M. Edelman, “Spike-timing dynamics of neuronal groups,” *Cerebral Cortex*, vol. 14, pp. 933–944, 2004.

E. M. Izhikevich, N. S. Desai, E. C. Walcott, and F. C. Hoppensteadt, "Bursts as a unit of neural information: Selective communication via resonance ,," *Trends Neurosci.*, vol. 26, pp. 161–167, 2003.

E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Networks*, vol. 14, pp. 1569–1572, Nov. 2003.

Latham, P. E., B. J. Richmond, P. G. Nelson, and S. Nirenberg. (2000). "Intrinsic Dynamics in Neuronal Networks. I. Theory.". *Journal of Neurophysiology* 88 (2): 808–27.

6 Neural Networks (1. Perceptrons and 2. Multi layered Perceptrons)

6.1 Perceptrons

In the previous chapter, by a progressive simplification of single neuron models, we arrived at the McCulloch-Pitts neuron model which takes inputs from many neurons and produces a single output. If the net effect of the external inputs is greater than a threshold, the neuron goes into excited state (1), else it remains in its resting state (0).

Using this model, in 1943, its inventors Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician, set out to construct a model of brain function. Note that it was the time of World War-II. It was also a time when use of computing power was being tested for the first time on a large scale for war purposes – for calculating missile trajectories and breaking enemy codes. The power of computing technology was just being realized by the world. Therefore it was natural to think of brain as a computer. Since the digital computer works on the basis of Boolean algebra, McCulloch and Pitts thought if it is possible for the brain also to use some form of Boolean algebra.

Since the MP neurons are binary units it seemed worthwhile to check if the basic logical operations can be performed by these neurons. McCulloch and Pitts quickly showed that the MP neuron can implement the basic logic gates AND, OR and NOT simply by proper choice of the weights:

OR Gate:

The truth table of an OR gate is:

X1	X2	Y
0	0	0
0	1	1
1	0	1
1	1	1

Note that the function below, which represents a MP neuron with two inputs, x_1 and x_2 , implements an OR gate.

$$y = g(x_1 + x_2 - b)$$

where $b = 0.5$; $g(\cdot)$ is the step function; $x_1, x_2 \in \{0, 1\}$. Actually any value of the bias term b , $0 < b < 1$, should work.

AND Gate:

The truth table of an AND gate is:

X1	X2	Y
0	0	0
0	1	0
1	0	0
1	1	1

Note that the function below implements an AND gate.

$$y = g(x_1 + x_2 - b)$$

where $b = 1.5$; $g(\cdot)$ is the step function; $x_1, x_2 \in \{0, 1\}$. Actually any value of the bias term b , $1 < b < 2$, should work.

NOT Gate:

$$y = g(-x + 0.5)$$

The truth table of a NOT gate is:

X	Y
0	1
1	0

Note that the function below implements a NOT gate.

$$y = g(-x + 0.5)$$

More generally, in $y = g(-x + b)$, any value of b in, $0 < b < 1$, would give a NOT gate.

Thus it became clear that by connecting properly designed MP neurons in specific architectures, any complex Boolean circuit can be constructed. Thus we have a theory of how brain can perform logical operations. McCulloch and Pitts explained their ideas in a paper titled, "A logical calculus of the ideas immanent in nervous activity" which appeared in the Bulletin of Mathematical Biophysics 5:115-133.

Although the idea of considering neurons as logic gates and the brain itself as a large Boolean circuit is quite tempting, it does not satisfy other important requirements of a good theory of the brain. There are some crucial differences between the brain and a digital computer (Table 6.1.1).

Table 6.1.1: Difference between the brain and a digital Computer

Property	Computer	Brain
Shape	2d Sheets of inorganic matter	3d volume of organic matter
Power	Powered by DC mains	Powered by ATP
Signal	Digital	pulsed
Clock	Centralized clock	No centralized clock
Clock speed	Gigahertz	100s of Hz
Fault tolerance	Highly fault-sensitive	Very fault-tolerant
Performance	By programming	By learning

Thus there are some fundamental differences between the computer and the brain. The signals used in the two systems are very different. There is no centralized clock in the brain. Each neuron fires at its own frequency which further changes with time. A brain is very fault tolerant which can be seen by the manner in which a stroke patient recovers. Most importantly a computer has to be programmed whereas the brain can learn by a progressive trial-and-error process.

These considerations led to the feeling that something is wrong with the McCulloch-Pitts approach to the brain.

As an answer to the above need, Frank Rosenblatt developed the Perceptron in 1957. A Perceptron is essentially a network of MP neurons.

[perceptron figure here – n inputs and m outputs]

Thus a Perceptron maps an m -dimensional input vector, onto a n -dimensional output vector. A distinct feature of a Perceptron is that the weights are not pre-calculated as in a MP neuron but are adjusted by a iterative process called training. The general approach to training, not only of a Perceptron, but of a larger class of neural networks (feedforward networks which will be defined later) is depicted in the figure below.

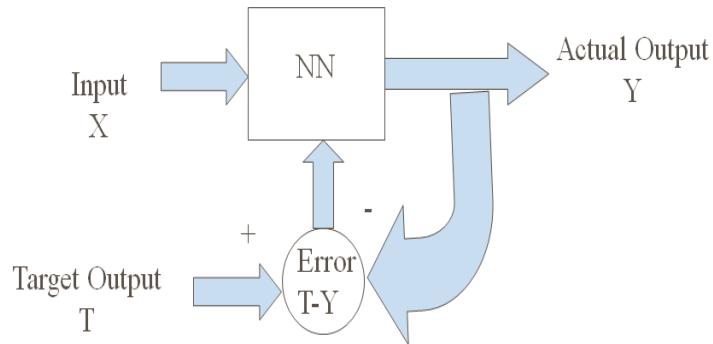


Figure 6.1.1: Training a neural network

The network is initialized with random weights.

When an input X is presented to a neural network (NN), it responds with an output vector Y . Since the weights are random, the network output is likely to be wrong and therefore different from a Desired or a Target output T . Error defined as $E = T - Y$, is used to adjust the weights in the Perceptron (or NN in general) in such a way that the next time when X is presented to the network, the response Y is likely to be closer to T than before. This iterative procedure is continued with a large number of patterns until the error is minimum on all patterns.

The mechanism by which the weights are adjusted as a function of the error is called the learning rule.

The learning rule can vary depending on the precise architectural details of the Neural Network (NN) used in the above scheme (Fig. 6.1.1).

Instead of directly taking up the task of deriving the learning rule for a Perceptron, let us begin with a very simple neuron model and derive the learning rule. In the process, we would introduce a few terms. The same procedure, with all its jargon, will be used to derive the learning rule for more complex architectures.

6.1.1 Case 1: Linear Neuron model: $y = \mathbf{w}^T \mathbf{x}$

Procedure to find the weights:

- 1) Noniterative, 2) Iterative

Output Error:

$$E = \frac{1}{2} \sum_n (y(p) - w^T x(p))^2$$

$$\nabla_w E = \sum_p (-x(p))(y(p) - w^T x(p)) = 0$$

1) Noniterative:

Pseudoinverse:

Let,

$$X = \begin{bmatrix} x(1) \\ \vdots \\ x(N) \end{bmatrix} = \begin{bmatrix} x_1(1) & x_m(1) \\ \vdots & \vdots \\ x_1(N) & x_m(N) \end{bmatrix}_{N \times m}$$

$$Y = X W.$$

$$w = [w_1 \dots w_m]^T$$

$$y = [y(1) \dots y(N)]^T$$

$$d = [d(1) \dots d(N)]^T$$

$$E = (1/2) (d - XW)^T (d - XW) = (1/2)d^T d - (X^T d)^T w + (1/2) w^T (X^T X) w \quad (6.1.1.1)$$

$$W = (X^T X)^{-1} X^T d$$

$$R_x = (X^T X)$$

correlation matrix

$$r_{xd} = X^T d$$

cross-correlation matrix

$$(X^T X)^{-1}$$

pseudo-inverse

2) Iterative:

a) Steepest Descent

$$w(t+1) = w(t) + \eta \sum_p (x(p))(d(p) - w^T x(p))$$

$$w(t+1) = w(t) + \eta(r_{xd} - R_x w(t)) \quad (6.1.1.2)$$

In the method of steepest descent, all the training data is used at the same time (packed into R_x and r_{xd}) to update every single weight. This can involve a large memory requirement and can be computationally expensive.

b) Least Mean Square Rule

$$w(t+1) = w(t) + (x(p))(d(p) - w^T x(p)) \quad (6.1.1.3)$$

Also called the Delta Rule, Widrow-Hoff Rule.

Note that the key difference between the steepest descent rule above (eqn. (6.1.1.2)) and the delta rule (eqn. (6.1.1.3)) is the absence of summation over all training patterns in the latter.

In this case the weight vector does not smoothly converge on the final solution. Instead, it performs a random walk around the final solution and converges only in a least square sense.

Issues:

1. Convergence

- 1a. Shape of Error function, E (Single minimum for quadratic Error function)

Note that the dominant term in the error function of eqn. (6.1.1.1) is a quadratic form associated with the correlation matrix, R_x .

Since R_x is a positive definite matrix the error function always has a unique minimum. It also has real, positive eigenvalues (λ_i).

- 1b. Effect of eigenvalues of correlation matrix

Condition number: $\lambda_{\max}/\lambda_{\min}$
 where λ_{\max} is the largest eigenvalue and λ_{\min} is the smallest eigenvalue of the correlation matrix R_x .

Slows down the descent over the error function if the condition number is too large

2. The need to choose η .

Large $\eta \rightarrow$ oscillations, instability

Small $\eta \rightarrow$ slow convergence

Bounds over the learning rate, η :

$$0 < \eta < 2/(\eta_{\max})$$

Proof:

$$R_x = (X^T X) \quad \text{correlation matrix}$$

$$r_{xd} = X^T d \quad \text{cross-correlation matrix}$$

$$(X^T X)^{-1} \quad \text{pseudo-inverse}$$

$$\begin{aligned} E &= (1/2) (d - XW)^T (d - XW) = (1/2)d^T d - (X^T d)^T w + (1/2) w^T (X^T X) w \\ &= (1/2) d^T d - r_{xd}^T w + (1/2) w^T (R_x) w \end{aligned}$$

$$\text{Final value of } w, w^* = R_x^{-1} r_{xd}$$

$$E = E_{\min} + (1/2)(w - w^*)^T R_x (w - w^*)$$

$$\text{Grad}(E) = R_x(w - w^*)$$

$$\begin{aligned} \Delta w &= -\eta \text{ grad}(E) \\ &= -\eta w(t+1) = w(t) - \eta R_x(w(t) - w^*) \\ w(t+1) - w^* &= (I - \eta R_x)(w(t) - w^*) \\ \text{Let, } v(t) &= (w(t) - w^*) \end{aligned}$$

$$v(t+1) = (I - \eta R_x) v(t) \quad (6.1.1.4)$$

Let $v' = Qv$, where Q is an orthogonal matrix that diagonalizes R_x .

$$v'(t+1) = (I - \eta D) v'(t) \quad (6.1.1.5)$$

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

If we consider the individual components, v_i , of eqn. (6.1.1.5) above,

$$v_i(t+1) = (1 - \eta \lambda_i) v_i(t) \quad (6.1.1.6)$$

The condition for stability of the last equation, is

$|1 - \eta \lambda_i| < 1$. Let us consider the two possible cases of this inequality.

a) $1 - \eta \lambda_i < 1 \Rightarrow \eta > 0$ which is trivial.

b) $-(1 - \eta \lambda_i) < 1 \Rightarrow \eta < 2 / \lambda_i$ for all i.

Therefore,

$$\eta < 2 / \lambda_{\max}$$

3. The need to reduce η with time:

There are obvious tradeoffs between use of a large vs. small η . Large η speeds up learning but can be unstable. Small η is stable but results in slower learning.

Therefore, it is desirable to begin with a large η and reduce it with time.

Learning rate variation schedules:

$$\text{a) } \eta = c/n; \quad \text{b) } \eta = \eta_0 / (1 + (n/\tau)) \quad (6.1.1.7)$$

6.1.2 Case 2: Perceptron: MP Neuron which has Sigmoid nonlinearity

With the hard-limiting or threshold nonlinearity the neuron acts as a classifier.

Final solution is not unique.

Convergence only if linearly separable.

$$y = g\left(\sum_{i=1}^n w_i x_i - b\right)$$

Hardlimiter characteristics:

$$g(v) = 1, v \geq 0$$

$$= 0, v < 0$$

For a Perceptron with a single output neuron, the regions corresponding to the 2 classes are separated by a hyperplane given by:

$$\sum_{i=1}^n w_i x_i - b = 0$$

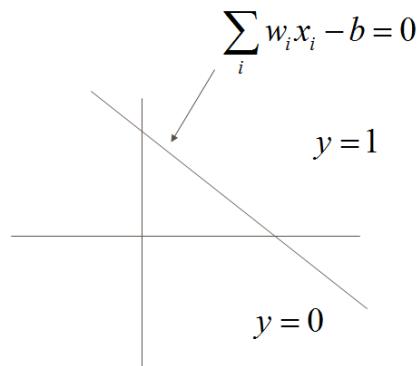


Figure 6.1.2.1: Classification by a perceptron

In other words, a Perceptron classifies input patterns by dividing the input space into two semi-infinite regions using a hyperplane.

6.1.3 Perceptron Learning Rule:

It is also called the LMS Rule or Delta Rule or Widrow-Hoff Rule

The steps involved in Perceptron learning are as follows:

1. Initialization of weights: Set the initial values of the weights to 0. $\mathbf{w}(0) = 0$.
2. Present the p'th training pattern, x , and calculate the network output, y .
3. Using the desired output, d , for the input pattern x , adjust the weights by a small amount using the following learning rule:

$$w(t+1) = w(t) + \eta[d(t) - y(t)]x(t)$$

where,

$$d(n) = +1, x(t) \in C1$$

$$d(n) = -1, x(t) \in C2$$

4. Go back to step 2 and continue until the network output error, $e = d-y$, is 0 for all patterns.

The above training process will converge after N_{\max} iterations where,

$$\begin{aligned} \alpha &= \min_{x(n) \in C1} w_0^T x(n) \\ \beta &= \max_{x(k) \in C1} \|x(k)\|^2 \\ N_{\max} &= \beta \|w_0\|^2 / \alpha \end{aligned} \quad (6.1.3.2)$$

See (Haykin 1999, Chapter 3, Section 3.9) for proof of convergence.

Range of η :

$$0 < \eta \leq 1$$

- Averaging of past inputs leads to stable weight dynamics, which requires small η
- Fast adaptation requires large η

Learning rule can also be derived from an error function:

$$E = \frac{1}{2} \sum_p [(d_p - y_p)^2] \quad (6.1.3.3)$$

where E denotes the squared error over all patterns.

The learning rule may be derived by performing gradient descent over the error function.

Gradient of Error:

$$\Delta w = -\eta \nabla_w E$$

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i}$$

$$\frac{\partial E}{\partial w_i} = -[d - y] \frac{\partial y}{\partial w_i} = -[d - y] g' x_i \quad (6.1.3.4)$$

The last term in the above equation has g' , which is zero everywhere except at the origin if g is a hardlimiting nonlinearity. But if we take a smoother version of $g()$, which saturate at +1 and -1, like the $\tanh()$ function, the learning rule becomes,

$$\Delta w_i = \eta [d - y] g' x_i \quad (6.1.3.5)$$

Since $g' > 0$ always for $\tanh()$ function, we can absorb it into h , considering it as a quantity that varies with x . We then have,

$$\Delta w_i = \eta [d - y] x_i \quad (6.1.3.6)$$

Which is identical to the Perceptron learning rule given in eqn. (6.1.3.1) above.

6.1.3.1 Features of Perceptron:

- 1) The Perceptrons can only classify linearly separable classes.

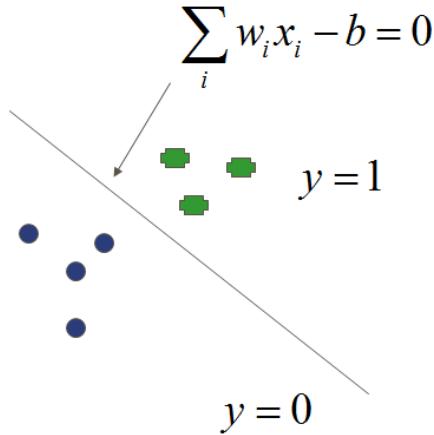


Figure 6.1.3.1.1: Classification of linearly separable classes by a perceptron

2. When the training data is linearly separable, there can be an infinite number of solutions.

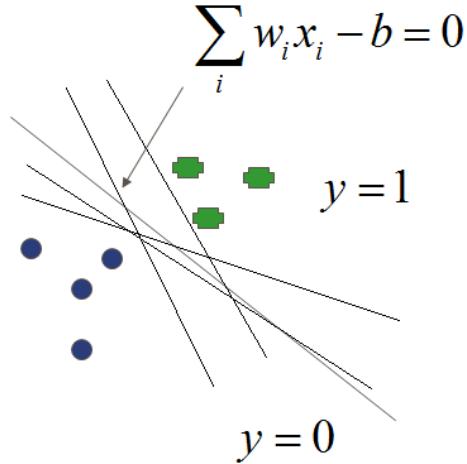


Figure 6.1.3.1.2: Solutions to classify linearly separable classes by a perceptron

6.1.3.2 Critique of Perceptrons:

- Perceptrons cannot even solve simple problems like Xor problem (Fig. 6.1.3.2.1)
- Linear model: Can only discriminate linearly separable classes
- Even the multi-layered versions may be afflicted by these weaknesses
(Minsky & Papert, 1969)

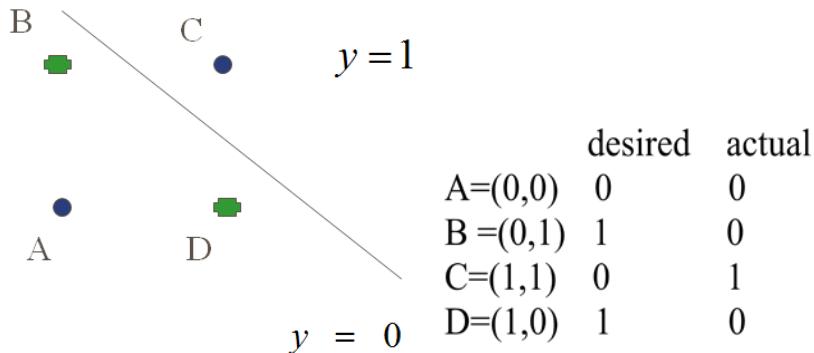


Figure 6.1.3.2.1: Inability to solve Xor by a perceptron

Exercises:

- 1) Perform gradient descent on,
 $E = x^2 + 10 y^2$, with h ranging from 0.01 to 0.2. What is the value of h at which instability occurs?
- 2) Generate the training set using the formula $f(x,y) = 2x + 3y + n$, where x and y are uniformly distributed over the interval [-1,1], and n is a Gaussian random variable with 0 mean and variance 0.5.
 - a) Train a 2-input, 1-output linear neuron model using the data set generated above. Find weights w₁ and w₂ (no bias) using all the 3 methods: i) pseudoinversion, ii) steepest descent and iii) LMS rule. Compare the 3 solutions. Comment on the variation of $w=(w_1, w_2)$ with time in case of the LMS rule.
- 3) Character recognition of digits (0-9) in 7-segment representation.
- 4) Represent digits (0 to 4) in a 10X10 array. Train a 100-input, 5 output perceptron to classify these 5 classes of characters. Generate several examples of each character by slightly shifting the character within the array, adding noise, randomly flipping some bits etc. Separate the data generated into training and test portions. Train the perceptron on the training data, freeze the weights and test it on the test data. Evaluate performance.

6.2 The Multi-layered Perceptron

Improvements over Perceptron:

- 1) Smooth nonlinearity - sigmoid
- 2) 1 or more hidden layers

6.2.1 Adding a hidden layer:

The perceptron, which has no hidden layers, can classify only linearly separable patterns. The MLP, with at least 1 hidden layer can classify *any* linearly non-separable classes also. An MLP can approximate any continuous multivariate function to any degree of accuracy, provided there are sufficiently many hidden neurons (Cybenko, 1988; Hornik et al, 1989). A more precise formulation is given below.

A serious limitation disappears suddenly by adding a single hidden layer.

It can easily be shown that the XOR problem which was not solvable by a Perceptron can be solved by a MLP with a single hidden layer containing two neurons.

XOR Example:

Neuron 1:

$$V_1 = \sigma(x_1 + x_2 - 1.5)$$

$$V_2 = \sigma(x_1 + x_2 - 0.5)$$

$$y = \sigma(V_1 - V_2 - 0.5)$$

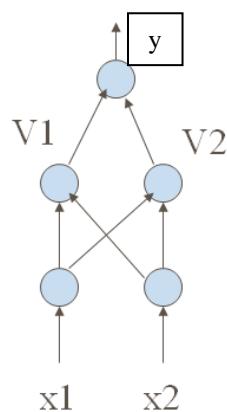


Figure 6.2.1.1: MLP for solving Xor

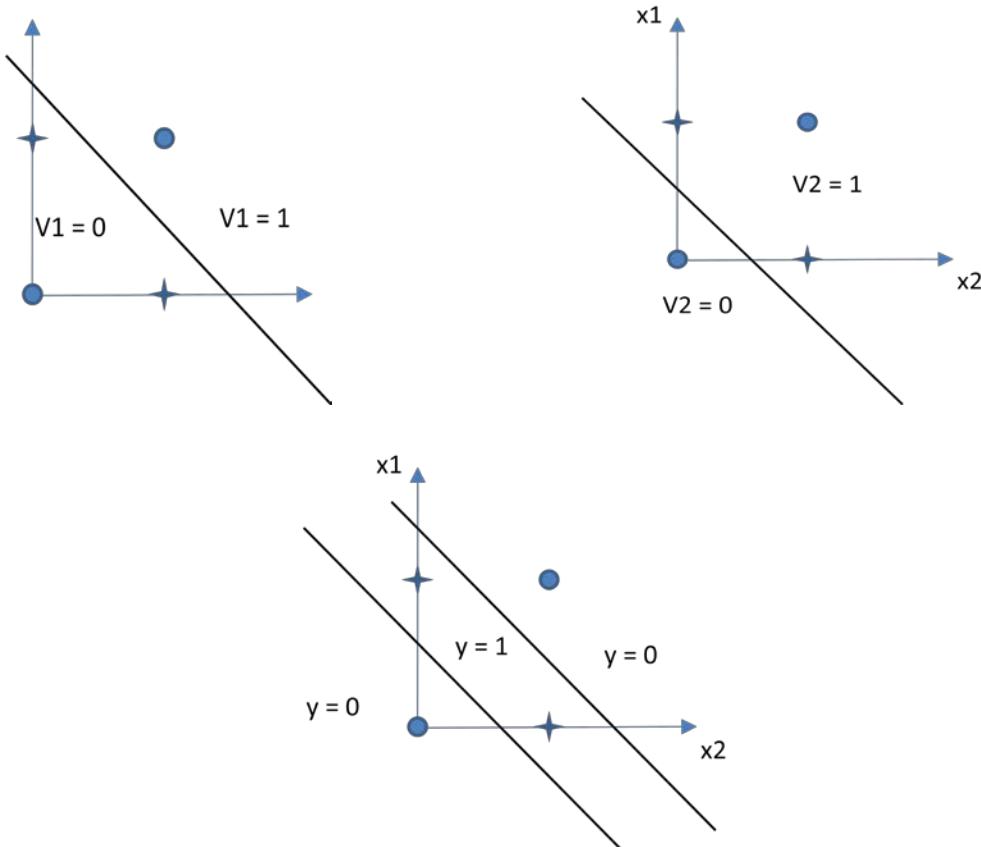


Figure 6.2.1.2: Plots showing the classification by $V1, V2$ and y (output of MLP)

6.2.2 Training the hidden layer:

Not obvious how to train the hidden layer parameters.

The error term is meaningful only to the weights connected to the output layer. How to adjust hidden layer connections so as to reduce output error? – *credit assignment* problem.

Any connection can be adapted by taking a full partial derivative over the error function, but then to update a single weight in the first stage we need information about distant neurons/connections close to the output layer (locality rule is violated). In a large network with many layers, this implies that information is exchanged over distant elements of the network though they are not directly connected. Such an algorithm may be mathematically valid, but is biologically unrealistic.

6.2.2.1 The Backpropagation Algorithm:

History:

1. First described by Paul Werbos (1974) in his PhD thesis at MIT.
2. Rediscovered by Rumelhart, McClelland and Williams (1986)
3. Also discovered by Parker (1985) and LeCun (1985) in the same year.

As in Perceptron, this training algorithm involves 2 passes:

- The forward pass – outputs of various layers are computed
- The backward pass – weight corrections are computed

Consider a simple 3-layer network with a single neuron in each layer.

$$\text{Total output error over all patterns: } E = \sum_p E_p \quad (6.2.2.1.1)$$

$$\text{Squared Output error for the } p\text{'th pattern: } E_p = \frac{1}{2} \sum_i e_i^2 \quad (6.2.2.1.2)$$

$$\text{Output error for the } p\text{'th pattern: } e_i = d - y_i \quad (6.2.2.1.3)$$

$$\text{Network output: } y_i = g(h_i^s) \quad (6.2.2.1.4)$$

$$\text{Net input to the output layer: } h_i^s = \sum_j w_{ij}^s V_j - \theta_i^s \quad (6.2.2.1.5)$$

$$\text{Output of the hidden layer: } V_j^f = g(h_j^f) \quad (6.2.2.1.6)$$

$$\text{Net input of the hidden layer: } h_j^f = \sum_k w_{jk}^f x_k - \theta_j^f \quad (6.2.2.1.7)$$

Update rule for the weights using gradient descent:

$$\Delta w_{ij}^s = -\eta \frac{\partial E_p}{\partial w_{ij}^s}; \quad \Delta \theta_i^s = -\eta \frac{\partial E_p}{\partial \theta_i^s} \quad (6.2.2.1.8)$$

$$\Delta w_{jk}^f = -\eta \frac{\partial E_p}{\partial w_{jk}^f}; \quad \Delta \theta_j^f = -\eta \frac{\partial E_p}{\partial \theta_j^f} \quad (6.2.2.1.9)$$

Updating w_{ij}^s :

$$\begin{aligned} \frac{\partial E_p}{\partial w_{ij}^s} &= \frac{\partial E_p}{\partial e_i} \frac{\partial e_i}{\partial y_i} \frac{\partial y_i}{\partial h_i^s} \frac{\partial h_i^s}{\partial w_{ij}^s} \\ &= e(-1)g'(h_i^s)V_j \end{aligned} \quad (6.2.2.1.10)$$

The delta at the output layer, δ_i^s , is defined as,

$$\delta_i^s = e \ g'(h_i^s) \quad (6.2.2.1.11)$$

Therefore,

$$\Delta w_{ij}^s = -\eta \frac{\partial E_p}{\partial w_{ij}^s} = \eta \delta_i^s V_j \quad (6.2.2.1.12)$$

By similar arguments, it can be easily be shown that the update rule for the threshold term is,

$$\Delta \theta_i^s = -\eta \delta_i^s \quad (6.2.2.1.13)$$

Updating w_{jk}^f :

$$\Delta w_{jk}^f = -\eta \frac{\partial E_p}{\partial w_{jk}^f}$$

$$\begin{aligned} \frac{\partial E_p}{\partial w_{jk}^f} &= \sum_i \frac{\partial E_p}{\partial e_i} \frac{\partial e_i}{\partial y_i} \frac{\partial y_i}{\partial h_i^s} \frac{\partial h_i^s}{\partial V_j} \frac{\partial V_j}{\partial w_{jk}^f} \\ &= \sum_i e_i (-1) g'(h_i^s) w_{ij}^s \frac{\partial V_j}{\partial w_{jk}^f} \\ &= \sum_i e_i (-1) g'(h_i^s) w_{ij}^s g'(h_j^f) x_k \\ &= \sum_i \delta_i^s w_{ij}^s g'(h_j^f) x_k \end{aligned} \quad (6.2.2.1.14)$$

Define an error term at the hidden layer as,

$$\delta_j^f = \sum_i \delta_i^s w_{ij}^s g'(h_j^f) \quad (6.2.2.1.15)$$

Therefore,

$$\Delta w_{jk}^f = \eta \delta_j^f x_k \quad (6.2.2.1.16)$$

Similarly the update rule for the threshold term is,

$$\Delta\theta_j^f = -\eta\delta_j^f \quad (6.2.2.1.17)$$

weight correction = (learning rate) * (local δ from ‘top’) * (activation from ‘bottom’)

General formulation of the Backpropagation Algorithm:

Notation:

Input (at k'th input neuron)	-	x_k
Actual Output (at i'th output neuron)	-	y_i
Target output (at i'th output neuron)	-	d_i
Hidden neuron activation	-	V_j^l (of j'th neuron in l'th layer)
Layer number	-	$l=0$ (input layer) to L (output layer)
Net input	-	h_{jl} (for j'th neuron in l'th layer)
$g(h)$	-	sigmoid nonlinearity $= 1/(1+\exp(-\beta h))$

Steps:

1. Initialize weights with small random values
2. (Loop over training data)
3. Choose a pattern and apply it to the input layer

$$V_k^o = x_k^p \quad \text{for all } k.$$

4. Propagate the signal forwards thro’ the network using:

$$V_j^l = g(h_j^l) = g(\sum_k w_{jk}^l V_k^{l-1} - b_j^l).$$

for each j, k and and ‘l’ until final outputs V_i^L have all been calculated.

5. Compute errors, δ ’s, for the output layer.

$$\delta_i^L = g'(h_i^L)[d_i(p) - V_i^L]$$

6. Compute δ ’s for preceding layers by backropagation of error:

$$\delta_i^{l-1} = g'(h_i^{l-1})[\sum_j w_{ji}^l \delta_j^l]$$

For $l = L, L-1, \dots, 1$

7. Update weights using the following:

$$\Delta w_{ij}^l = \eta \delta_i^l V_j^{l-1};$$

$$\Delta \theta_i^l = -\eta \delta_i^l$$

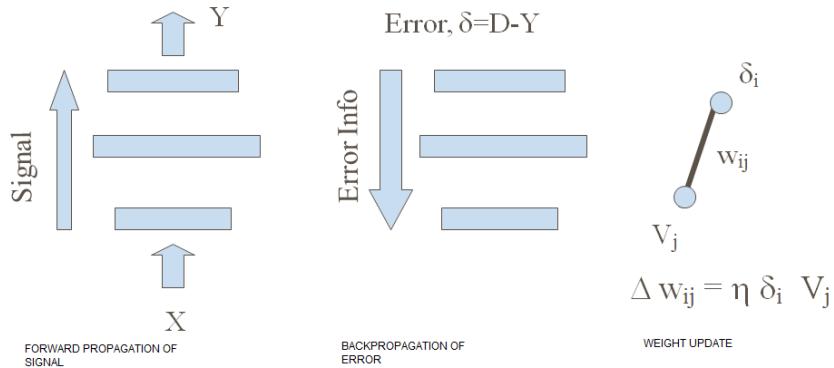


Figure 6.2.2.1: Training using Back propagation algorithm

Traning:

- Randomly initialize weights.
- Train network using backprop eqns.
- Stop training when error is sufficiently low and freeze the weights.

Testing

- Start using the network.

Merits of MLP trained by BP:

- a) A general solution to a large class of problems.
- b) With sufficient number of hidden layer nodes, MLP can approximate arbitrary target functions.
- c) Backprop applies for arbitrary number of layers, partial connectivity (no loops).
- d) Training is local both in time and space – parallel implementation made easy.
- e) Hidden units act as “feature detectors.”
- f) Good when no model is available

Problems with MLP trained by BP:

- a) Blackbox approach
- b) Limits of generalization not clear
- c) Hard to incorporate prior knowledge of the model into the network
- d) slow training
- e) local minima

6.2.3 Architecture of MLP:

If there is no nonlinearity then an MLP can be reduced to a linear neuron.

1. Universal Approximator:

Theorem:

Let $g(\cdot)$ be a nonconstant, bounded, and monotone-increasing continuous function. Let I_m denote the m -dimensional hypercube $[0,1]^m$. The space of continuous functions on I_m is denoted by $C(I_m)$. Then given any function f in $C(I_m)$ and $\epsilon > 0$, there exist an integer M and sets of real constants a_i, b_i and w_{ij} , where $i = 1, \dots, n$, and $j = 1, \dots, m$ such that we may define:

$$F(x_1, \dots, x_m) = \sum_{i=1}^n a_i g\left(\sum_{j=1}^m w_{ij} x_j + b_i\right)$$

As an approximate realization of the function $f(\cdot)$; that is,

$$|F(x_1, \dots, x_m) - f(x_1, \dots, x_m)| < \epsilon$$

for all x_1, \dots, x_m that lie in the input space.

For the above theorem to be valid, the sigmoid function $g(\cdot)$ has to satisfy some conditions. It must be: 1) non-constant, 2) bounded, 3) monotone-increasing and 4) continuous.

All the four transfer functions described in the section on Perceptrons satisfy conditions #1,2 and 3. But the hardlimiting nonlinearities are not continuous. Therefore, the logistic function or the tanh function are suitable for use as sigmoids in MLPs.

2. In general more layers/nodes greater network complexity

2.1 Although 3 hidden layers with full connectivity are enough to learn any function often more hidden layers and/or special architectures are used.

More hidden layers and/or hidden nodes:

3-layer network:

arbitrary continuous function over a finite domain

4-layer network

Neurons in a 3-layer architecture tend to interact globally.

In a complex situation it is hard to improve the approximation at one point without worsening it at another.

So in a 4-layer architecture:

1st hidden layer nodes are combined to construct locally sensitive neurons in the second hidden layer.

Discontinuous functions:

learns discontinuous (inverse function of continuous function) functions also (Sontag, 1992)

For hard-limiting threshold functions:

- 1st hidden layer: semi-infinite regions separated by a hyper-plane
- 2nd hidden layer: convex regions
- 3rd hidden layer: non-convex regions also

6.2.4 Training MLP:

1. Initialization: is VERY important.

$g'(\cdot)$ appears on the right side of all weight update rules (Refer sections 6.1.1, 6.1.2, 6.2.1). Note that $g'(\cdot)$ is high at the origin and falls on both sides. Therefore most learning happens when the net input (h) to the neurons is close to 0. Hence it is desirable to make initial weights small. A general rule for initialization of input weights for a given neuron is:

$$\text{Mean}(w(0)) = 0.$$

$$std(w(0)) = \frac{1}{\sqrt{m}} \text{ where } m \text{ is the number of inputs going into a neuron.}$$

2. Batch mode and Sequential mode:

Epoch: presentation of all training patterns is called an epoch.

Batch mode:

Updating network weights once every epoch is called batch mode update.

- memory intensive
- greater chance of getting stuck in local minima

Sequential mode:

Updating the network weights after every presentation of a data point is sequential mode of update.

- lesser memory requirement
- The random order of presentation of input patterns acts as a noise source lesser chance of local minima

Rate of learning:

We have already seen the tradeoffs involved in choice of a learning rate.

Small learning rate \rightarrow approximate original continuous domain equations more closely but slows down learning.

Large learning rate $\eta \rightarrow$ poorer approximation of original equations. Error may not decrease monotonically and may even oscillate. But learning is faster.

A good thumb rule for choosing eta ' η ':

$$\eta = 1/m$$

Where 'm' is the number of inputs to a neuron. This rule assumes that there are different η 's for different neurons.

3. Important tip relating learning rate and error surface:

Rough error surface → slow down, low eta

Smooth (flat) error surface → speed up, high eta

- i) Momentum:

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n)$$

Action of momentum:

$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^n \alpha^{n-t} \delta_j(t) y_i(t) = -\eta \sum_{t=0}^n \alpha^{n-t} \frac{\partial E(t)}{\partial w_{ji}(t)}$$

- a) If $|\alpha| < 1$, the above time-series is convergent.
- b) If the sign of the gradient remains the same over consecutive iterations the weighted sum delta w_{ji} grows exponentially i.e., accelerate when the terrain is clear.
- c) If the gradient changes sign in consecutive iterations, delta w_{ji} shrinks in magnitude i.e., slow down when the terrain is rough.

- ii) Separate eta for each weight:

- a) Separate η for each weight
- b) Every eta varies with time
- c) If $\Delta(w)$ changes sign several times in the past few iters, decrease η
- d) If $\Delta(w)$ doesn't change sign in the past few iters, increase η

Stopping Criteria: when do we stop training?

- a) Error < a minimum.
- b) Rate of change in error averaged over an epoch < a minimum.
- c) Magnitude of gradient $\|g(w)\| < a$ minimum.
- d) When performance over a test set has peaked.

Premature Saturation:

All the weight modification activity happens only when $|h|$ is within certain limits.

$g'(h) \approx 0$, or $\Delta(w) = 0$, for large $|h|$.

NN gets stuck in a shallow local minimum.

Solutions:

- 1) - Keep a copy of weights
- Retract to pre-saturation state
- Perturb weights, decrease η and proceed
- 2) - Reduce sigmoid gain (λ) initially
 - e) Increase λ gradually as error is minimized
 - f) $\lambda \rightarrow 0$ means the NN is a linear model in the operating range.
 - g) So there is only one minimum
- 3) $g'(h) \leftarrow g'(h) + \epsilon$
Quick prop
Derivative is never 0.
Network doesn't get stuck, but never settles either.
Again $\epsilon \rightarrow 0$ as the network approaches a minimum.

6.2.5 Testing/generalization:

Idea of overfitting or overtraining:

Using too many hidden nodes, may cause overtraining. The network might just learn noise and generalize poorly.

Example of polynomial interpolation:

Consider a data set generated from a quadratic function with noise added. A linear fit is likely to give a large error. Best fit is obtained with a quadratic function. Fit 10th degree might give a low error but is likely to learn the variations due to noise also. Such a fit is likely to do poorly on a test data set. This is called overfitting or poor generalization.

This happens because there are many ways of generalizing from a given training data set.

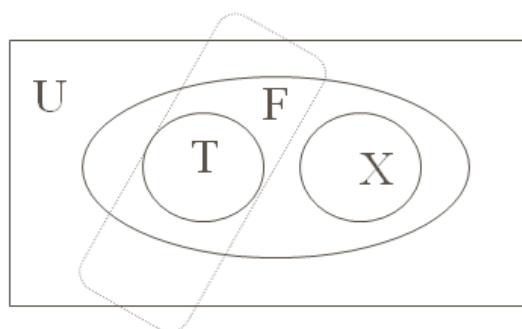


Figure 6.2.5.1: Existence of a multitude of ways of generalization

The above Venn diagram illustrates the possibility of generalizing in multiple ways from a given training data set. U is the universe of all possible input-output patterns. F (the ellipse) represents

the set of I/O pairs that define the function to be learnt by the mlp. T (circle) denotes the training data set which is a subset of F. X denotes the test data set. The dotted rectangle denotes the actual function learnt by the NN, which is consistent with the training set T, but is completely non-overlapping with the test set X, and very different from the unknown function F.

A simple calculation from (Hertz et al 1991).

Boolean function

N-inputs, 1-output

2^N patterns and 2^{2^N} rules totally.

Assume,

p – training patterns say, represent the rule T.

Then there are $(2^N - p)$ test patterns and there are $2^{(2^N - p)}$ rules, R, consistent with rule T.

$N = 30, p = 1000$ patterns.

You have 2^{10^9} generalizations exist for the same training set T.

6.2.6 Applications of MLP

Three applications of MLPs that simulate aspects of sensory, motor or cognitive functions are described.

1. Nettalk
2. Past tense learning
3. Autonomous Land Vehicle in a Neural Network (ALVINN)

6.2.6.1 NetTalk: A neural network that can read text aloud (Sejnowski and Rosenberg 1986)

Nettalk is a system that can read English text aloud and can pronounce letters accurately based on context (Sejnowski and Rosenberg 1986). It uses a three layer MLP.

Background:

English is not a phonetic language.

Char → Sound mapping is not unique.

The same character is pronounced differently depending on the context.

Examples:

The character ‘c’ is pronounced as /k/ in “cat” and as /s/ in “façade.”

The letters “-ave” form a long vowel in “gave” and “brave” but not in “have.”

Similarly, the letters “-ea-“ are pronounced as /ii/ (long vowel) in “read” (present tense) (pronounced as “reed”) and as /i/ (short vowel) as in “read” (past tense) (pronounced as “red”).

6.2.6.1.1 Network Architecture

Input representation:

26 alphabets and three punctuation marks (comma, fullstop and blank space) are supported. Input characters are not presented one at a time to the network. In order to incorporate the context, each character is presented along with a context which consists of three additional characters on either side of the character. Thus text is presented as windows of 7 symbols. Therefore the number of neurons to the input layer are:

$$(26+3) \times 7 \text{ input neurons}$$

Hidden layer : 80 hidden neurons

Output representation :

26 output neurons encoding phonemes.

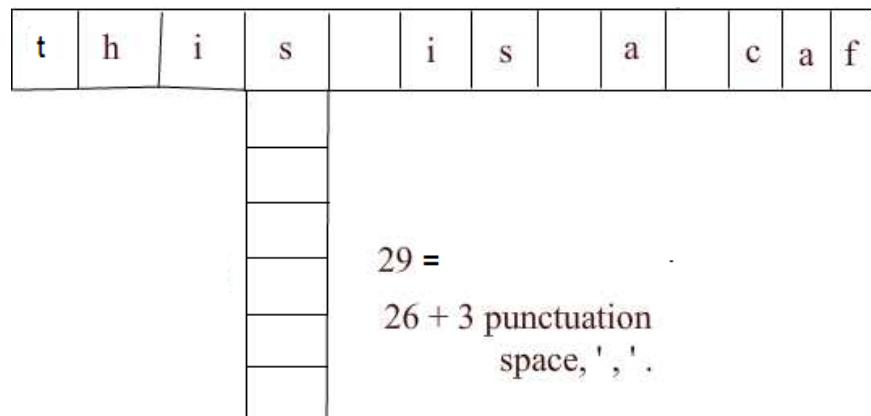


Figure 6.2.6.1.1: Input Representation

Network architecture: Ref: (fig. 6.8 in Hetz et al 1991)

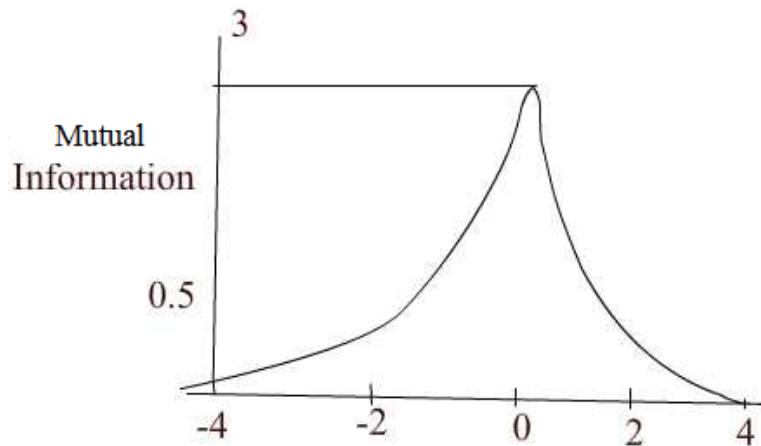


Figure 6.2.6.1.2: Mutuality in the window

A context of window of 7 characters is chosen since mutual information from a central character falls off rapidly in both directions as shown in the above figure. Beyond 3 neighbors mutual information falls to less than 0.1.

6.2.6.1.2 Training data

- Phonetic transcription from informal continuous speech of a child
- 20,000 words from a pocket dictionary

Eg: phone

$\Rightarrow f - o n -$

A set of 1000 words were chosen from this dictionary; these are selected from the Brown corpus of the most frequent words in English.

6.2.6.1.3 Learning curve

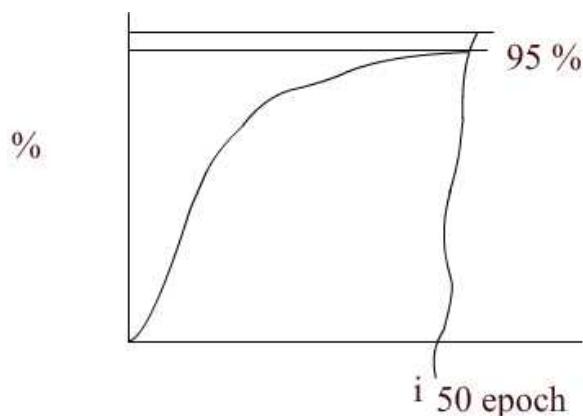


Figure 6.2.6.1.3: Performance curve

Found in human skill learning

Performance reaching about 95% correct after training the network with about 50,000 words.

The presence of a hidden layer is found to be crucial to achieve this high level of performance. When a two-layer network (perceptron) was used for the same problem, performance quickly rose to 82% and saturated at that level.

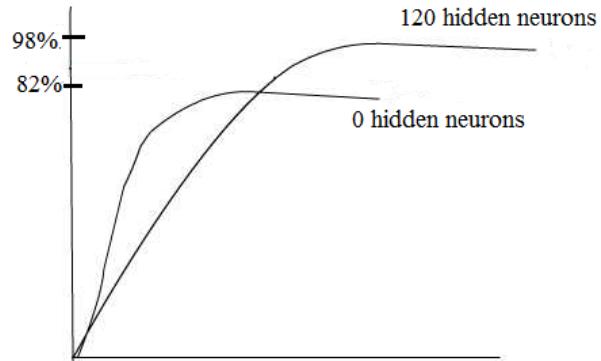


Figure 6.2.6.1.4: Performance curve on alteration of the hidden neurons

6.2.6.1.4 Stages of learning

1. One of the first features learnt by the network is distinction between vowels and consonants
2. Next the network learnt to pause at word boundaries. Output resembled pseudo words with pauses between them.
3. After 10 passes text was understandable

Error patterns

- Errors were meaningful
 - e.g thesis
 - these
- The network rarely confused between vowels & consonants
- Some errors actually were due to errors in annotation.

6.2.6.1.5 Test Performance

439 words from the regular speech. These words were not present in the training set.

Performance was at 78%. Thus the network demonstrated that it can generalize from one set of words to another.

Damage testing

In order to test the robustness of the network in face of damage, random noise, n , uniformly distributed between -0.5 and 0.5 was added to each weight in the trained network.

$$\omega_i \leftarrow \omega_i \pm v$$

$$v \in [-0.5, 0.5]$$

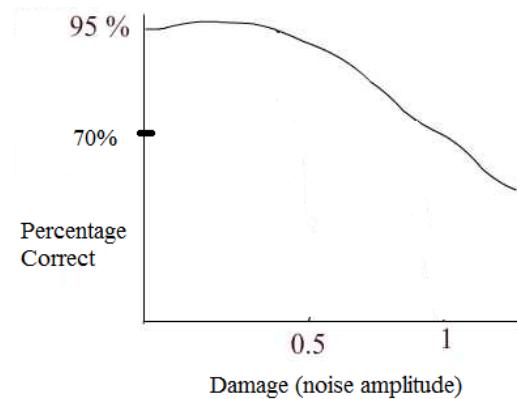


Figure 6.2.6.1.5: Performance on damage

The network was retrained after damage. The damaged network showed 67% performance. Therefore to compare the learning rate of the damaged network with the original network, the original learning curve was considered from a performance level of 67%. Figure 6.2.6.1.6 below shows that although the damaged network begins with the same performance level as the original network, it relearns the task much faster than the original network.

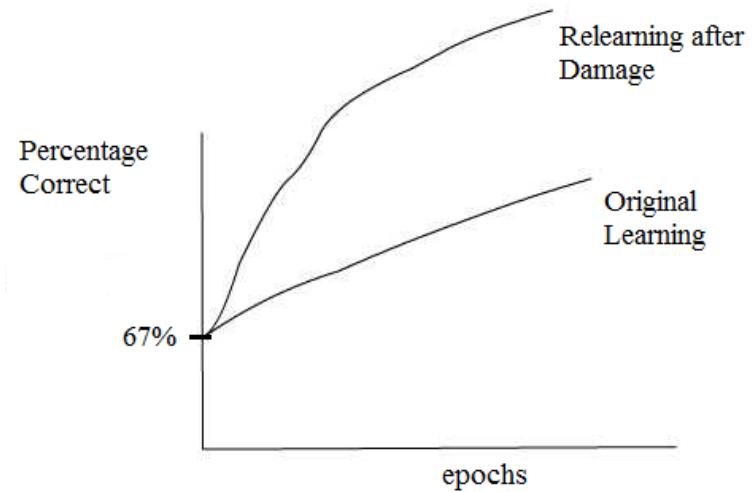


Figure 6.2.6.1.6: Relearning after damage

A careful analysis of the performance of the network specific letter-to-phoneme correspondences showed that different letter-to-phoneme correspondences were learnt at different rates.

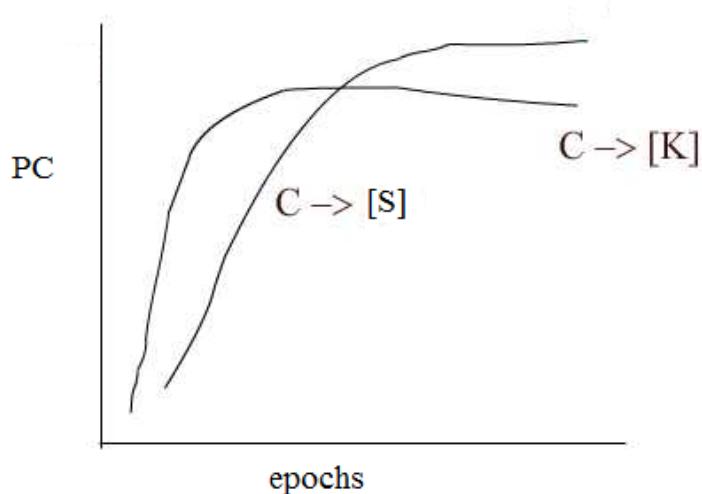


Figure 6.2.6.1.7: Rates of learning for different letter to phoneme representation

Soft ‘c’ (as in ‘nice’) takes longer to learn than hard ‘c’ (as in ‘cat’). This is probably because hard c occurs about twice as often as soft c in the corpus.

Interestingly children show a similar difficulty in learning.

Ref: Sejnowski, TJ Rosenberg, CR (1986). NETtalk: A parallel network that learns to read aloud, Tech. Rep. No. JHU/EECS-86/Q1

6.2.6.2 Past tense learning (Rumelhart & McClelland, 1986)

Three stages of past tense learning is seen in children.

Stage 1:- Only a small number of words correctly used correctly in past tense

- High frequency words, majority are irregular
- Children tend to get the past tense quickly.
- Typical examples: Came, got, gave, looked, needed, took, went.

Stage 2:-

- Children use much larger number of words
- Many verbs are used with correct past tense forms
Majority are regular
Eg. Wiped, pulled
- Children now incorrectly apply regular past tense endings for words which they used in stage 1

Eg: come ⇒ comed or camed

Stage 3:- Regular or irregular forms exist.

- Children have required the use of correct irregular form of past tense
- But continue to apply the regular form to new words they learn.

6.2.6.2.1 Training

Words

10- high frequency verbs (8 irregular + 2 regular)
Came, get, give, look, take, go, have, live, feel

410- medium frequency verbs,
334 regular, 76 irregular

86-low frequency words,
72 regular, 14 irregular

Stage 1:-

10 epochs of high frequency verbs
Enough to produce good performance

Stage 2:-

410 medium frequency verbs were added to 10 verbs (trained for 190 more epochs after phase 2). In this stage errors suddenly start creeping in. Most of the errors are due to regularization of irregular verbs (see fig. 8 below).

Stage 3:-

86 low frequency verbs are tested without training. Beyond state 2, the performance over regulars and irregulars is nearly the same, each touching 100%.

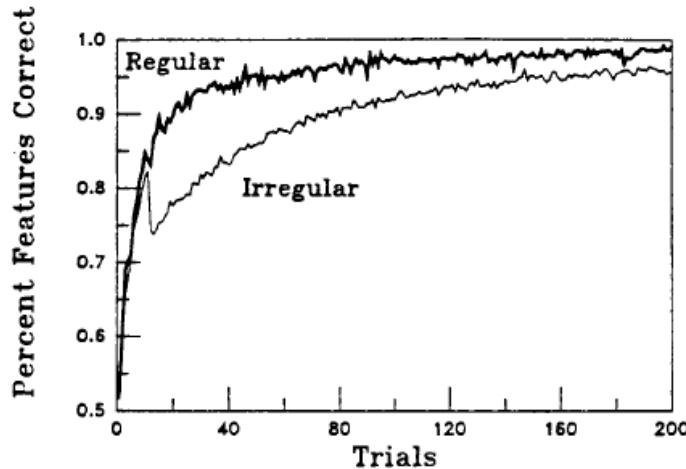


Figure 6.2.6.2.1: Performance curve on past tense learning

Ref: Rumelhart and McClelland, On learning the past tenses of English verbs, Parallel Distributed Processing, Vol. II, 1986

6.2.6.3 Autonomous Land Vehicle in a Neural Network (ALVINN)

6.2.6.3.1 Training:-

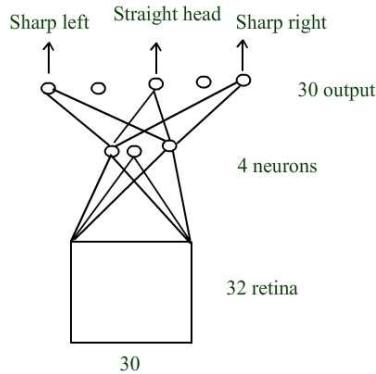


Figure 6.2.6.3.1: Network Architecture - ALVINN

$\frac{1}{4}$ to $\frac{1}{2}$ mile of training road which takes about 2 minutes to cover.

50 real images

Transformed 15 times to create 750 images.

6.2.6.3.2 Generalization:

Generalized to new points of the road

Different weather & lightning conditions

6.2.6.3.3 Specialization

Separate networks are trained on:

- Single-lane dirt & paired roads
- Two-lane suburban roads/city, streets
- multilane divided (90 mph) highways

6.2.6.3.4 Hidden layer Analysis

Hidden layer

Unlined paved roads → “edge detectors” trapezoidal shaped road regions

Lined highways → detectors for line markings

Dirt roads → ‘rut’ detectors.

Reference:

Dean Pomerleau, "ALVINN: An Autonomous Land Vehicle In a Neural Network,"
Advances in Neural Information Processing Systems 1, 1989.